Nancy-Université
Université Henri Poincaré

Département de formation doctorale en informatique **École doctorale IAEM Lorraine**
Sciences et Technologies

Modélisation stochastique pour le raisonnement médical et ses applications à la télémédecine

THÈSE

présentée et soutenue publiquement le 27 mai 2011

pour l'obtention du

Doctorat de l'université Henri Poincaré – Nancy 1

(spécialité informatique)

par

Cédric ROSE

Composition du jury

Président :	Christian Roux	Professeur à l'ENST Bretagne
Rapporteurs :	François Brémond	Directeur de recherche, INRIA
	Paul Rubel	Professeur à l'INSA de Lyon
Examinateurs :	Jacques Duchêne	Professeur à l'Université de technologie de Troyes
	Jacques Demongeot	Professeur à l'Université Joseph Fourier de Grenoble
	Jean-Paul Haton	Professeur à l'Université Henri Poincaré - Nancy 1
	Michèle Kessler	Professeur à l'Université Henri Poincaré - Nancy 1
Directeur de thèse :	François Charpillet	Directeur de recherche, INRIA

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Loria

Mis en page avec la classe thloria.

Remerciements

Je tiens à remercier tout ceux qui ont contribué, directement ou indirectement, à faire de cette thèse une expérience scientifique et humaine particulièrement enrichissante.

Je remercie en particulier François Charpillet, mon directeur, pour son encadrement mais aussi et surtout pour m'avoir convaincu de faire cette thèse.

Je souhaite également remercier Luis Vega, directeur de la société Diatélic, pour avoir accepté de financer ce travail et pour la liberté qu'il m'a donné pendant ces années de thèse.

Je remercie également l'ensemble des membres de mon jury :
- Christian Roux, Professeur à l'ENST Bretagne, qui m'a fait l'honneur de présider ce jury de thèse,
- François Brémond, Directeur de recherche à l'INRIA, et Paul Rubel, Professeur à l'INSA de Lyon, pour avoir accepté d'être rapporteurs de ce travail et pour leur remarques qui ont contribué à l'amélioration de ce mémoire,
- Jacques Duchêne, Professeur à l'Université de technologie de Troyes, Jacques Demongeot, Professeur à l'Université Joseph Fourier de Grenoble, Jean-Paul Haton, Professeur à l'Université Henri Poincaré de Nancy et Michèle Kessler, Professeur à l'Université Henri Poincaré de Nancy pour avoir accepté de faire partie du jury et pour leurs remarques et leurs questions constructives.

Cette thèse n'aurait pas pu se faire sans la participation active d'experts humains qui ont, en quelque sorte, été les sujets de la modélisation. Je remercie tout particulièrement Jacques Chanliau, néphrologue et directeur de l'ALTIR pour ses précieuses explications, toujours très claires, sur l'insuffisance rénale et sur la dialyse, pour sa disponibilité et pour son implication dans la conception du système expert pour le suivi de l'hydratation en hémodialyse. Je tiens également à remercier le Professeur Luc Frimat, chef du service de néphrologie au CHU de Nancy, pour ses explications sur la transplantation rénale et pour son implication dans la conception d'un système expert pour le suivi de patient transplantés. Je remercie aussi Bernard Béné, médecin et expert de la Société Gambro pour ses explications sur l'abord vasculaire en hémodialyse.

Je souhaite également remercier les collègues et amis qui ont contribué directement à ce travail, en particulier Jamal Saboune, Julien Thomas, Abdallah Dib, Amandine Dubois, Yoann Bertrand-Pierron, Renato Da Costa et Marion Casolari.

Je remercie également Olivier Buffet et Sylvain Castagnos avec une pensée particulière pour la préparation de ma soutenance.

Je remercie tout particulièrement mon père, pour son soutien et pour ses relectures.

Enfin je remercie tous les collègues et amis du Loria et toute l'équipe de Diatélic qui pendant ces années de thèse ont contribué par leur bonne humeur à faire de ce travail une expérience extrêmement positive que je recommande à tous ceux qui hésiteraient encore à faire une thèse.

Table des matières

Table des figures

Liste des tableaux

Liste des algorithmes

Introduction

Ce travail de thèse a été financé dans le cadre d'une convention CIFRE entre l'Institut National de Recherche en Informatique et Automatique (INRIA) et la société Diatélic spécialisée dans le développement d'applications de télémédecine. Il s'articule autour de plusieurs projets de recherche et développement menés en partenariat avec le CHRU de Nancy, l'Association Lorraine pour le Traitement de l'Insuffisance Rénale (ALTIR) et la société Gambro.

Nous débutons ce mémoire par la présentation du cadre applicatif qu'est la télémédecine, au sens large, en essayant d'en présenter les motivations et les enjeux. Les chapitres suivants sont consacrés à l'étude de ce que l'intelligence artificielle et, en particulier, la modélisation stochastique peuvent apporter à ce domaine.

1 Télémédecine

La télémédecine est un domaine relativement nouveau de la médecine reposant sur l'utilisation des technologies de l'information et des communications pour la pratique à distance de la médecine. C'est un terme générique qui recouvre des notions comme :

- la télésurveillance qui consiste en l'analyse régulière de données transmises en vue de la génération d'alertes,
- la téléconsultation qui permet à un praticien d'interagir à distance avec son patient,
- la téléexpertise qui désigne la sollicitation de l'avis d'un expert distant,
- la téléchirurgie qui va de l'opération assistée par ordinateur où le chirurgien n'est pas en contact direct avec son patient à l'opération à grande distance, comme dans le cas de l'« Opération Lindbergh », une opération chirurgicale réalisée en 2001 au CHU de Strasbourg par un chirurgien se trouvant à New York.

Au delà des défis scientifiques et technologiques, le développement de la télémédecine a pour ambition d'améliorer la qualité des soins en proposant de nouveaux outils et une nouvelle organisation des soins.

1.1 Motivations

L'une des ambitions suscitant le développement de la télémédecine est de répondre à un enjeu démographique. L'Observatoire National de la Démographie des Professions de Santé (ONDPS) indique dans son rapport annuel 2006-2007 que, si les effectifs de l'ensemble des professions de santé ont augmenté d'un peu plus de 20% entre 2000 et 2007, cette augmentation est plus modérée chez les médecins (7%). En 2025, le nombre de médecins devrait chuter de 10% en raison des départs à la retraite alors que le vieillissement de la population devrait s'accompagner d'un accroissement des besoins de soins. Le rapport de l'ONDPS souligne de grandes disparités géographiques dans l'évolution de la démographie médicale. Ainsi le Centre et la Picardie sont classés parmi les régions les plus désertées médicalement à l'opposé de la région PACA où l'on ne compte pas de commune à faible densité médicale. Il existe également de fortes disparités entre spécialités médicales. Dans ce contexte, la télémédecine peut être vue comme un moyen d'équilibrage permettant l'interaction à distance entre le patient et son équipe médicale, et entre les professionnels de santé eux même, favorisant ainsi une organisation des soins moins cloisonnée géographiquement.

L'enjeu est aussi celui de la prévention avec bien évidemment des aspects financiers. Nous pouvons ainsi citer le problème de la iatrogénie médicamenteuse. Selon une étude menée en

1998 par l'agence du médicament, les effets indésirables des médicaments constituent le motif d'admission d'un peu plus de 3% des hospitalisations ce qui représenterait 128 000 hospitalisations par an en France. Le rapport mentionne que, dans 31% des cas, ces effets indésirables relèvent de pratiques de prescription par le médecin ou d'utilisation par le patient non conformes à l'autorisation de mise sur le marché. Sur un plan comptable, cela représente 355 000 journées d'hospitalisations par an dont une grande partie pourrait être évitée. Dans une conférence donnée en 2006 au Collège de France, le professeur Patrice Degoulet de l'hôpital Georges-Pompidou relevait les points suivants :

- selon un rapport de l'institute of medicine aux Etats Unis (2000, 2001), la iatrogénie médicamenteuse serait à l'origine de 3 à 4 % des séjours hospitaliers avec une mortalité qui, extrapolée à la population américaine totale, serait supérieure à celle des accidents de la route.
- 50 % des décès par iatrogénie médicamenteuse pourraient être évités par l'utilisation d'un dossier électronique accompagné d'une aide à la prescription.

Plus généralement, l'enjeu est celui de l'amélioration de la prise en charge des malades. Comme le montrent diverses expériences de télésurveillance médicale, parmi lesquelles celle de la société Diatélic dans le cadre du suivi de l'insuffisance rénale, la prévention des complications chez les patients atteints de maladies chroniques peut être améliorée par la mise ne place d'un suivi régulier. Le rapport établi en 2008 par Pierre Simon et Dominique Acker sur la place de la télémédecine dans l'organisation des soins [Simon and Acker, 2008] cite ainsi différents domaines où la télésurveillance médicale a démontré un intérêt réel pour le patient comme la télésurveillance de l'hypertension artérielle, des grossesses à risques, des maladies cardiaques ou respiratoires, de l'insuffisance rénale ou encore du diabète. Le rapport relève que le nombre de patients atteints de maladies chroniques ne cesse d'augmenter avec l'allongement de la durée de la vie : 15 millions de patients sont atteints de maladies chroniques, aujourd'hui en France, et le chiffre annoncé est de 20 millions de patients pour 2020.

L'aide au maintien à domicile pour les personnes en perte d'autonomie, la détection et même la prévention des chutes ou encore la mesure de l'activité au domicile sont autant de domaines où la télésurveillance médicale a un rôle à jouer en raison de l'augmentation du nombre de personnes concernées.

1.2 Comment ?

Les apports de la télémédecine sont multiples et les systèmes d'aide à la décision peuvent avoir différents fondements.

L'aide à la décision repose avant tout sur la capacité du système à proposer une information pertinente. En rendant possible une meilleure circulation de l'information les T.I.C. (Technologies de l'Information et des Communications) nous donnent les moyens de développer des systèmes d'alertes visant à anticiper les complications et à aider le médecin dans sa prise de décision.

Le système peut être fondé sur des protocoles afin d'aider au respect des bonnes pratiques de prescription qu'il s'agisse de traitements ou d'examens. Le système peut apporter de l'information sur le degré de normalité des examens. D'un point de vu plus général, sa fonction est alors de rendre disponible une base de connaissances adaptée aux besoins du professionnel de santé.

L'aide à la décision peut se référer à des systèmes proposant de nouvelles sources de données. Le développement de capteurs portés par le patient ou équipant son domicile permet de proposer

un suivi en milieu écologique [Paysant, 2006], c'est à dire du sujet dans son environnement habituel. Le médecin peut alors disposer d'une information clinique beaucoup plus riche.

Le développement d'une intelligence collective est rendu possible par le partage de l'information au sein d'un réseau de professionnels de santé. Le système est alors le support des interactions entre le ou les spécialistes, le médecin traitant, les infirmières et, bien évidement, le patient qui est au centre du dispositif.

Ces systèmes participent également à la création de nouvelles connaissances. L'information recueillie aujourd'hui pourra être demain revisitée dans le cadre d'études scientifiques. Ils permettent aussi une meilleure évaluation des pratiques thérapeutiques. Le suivi en milieu écologique donne une information nouvelle qu'il faut apprendre à interpréter.

Ces systèmes de télémédecine reposent sur une informatisation au préalable des données à traiter.

1.3 L'informatisation des données médicales

Il existe actuellement de nombreuses expériences de télémédecine indépendantes les unes des autres. L'un des enjeux est l'interopérabilité des systèmes et des services. La notion d'hébergeur de données de santé est précisée dans la loi du 4 mars 2002 relative aux droits des malades et à la qualité du système de santé et le Dossier Médical Personnel (DMP) a été créé par la loi du 13 aout 2004. Il s'agit d'un « porte documents » électronique accessible par internet contenant l'ensemble du dossier médical. Après une première expérimentation en 2006 et un démarrage raté en 2007, le DMP a fait l'objet d'un second appel d'offre en 2009. Il est prévu que l'utilisation du DMP ne soit pas obligatoire mais, en revanche, les remboursements de l'assurance maladie seraient, en partie, conditionnés à l'utilisation du DMP ce qui constituerait une incitation certaine. Il est à prévoir que le développement du DMP puisse jouer un rôle central dans le développement et la coordination des systèmes de télémédecine. En 2009, la création de l'Agence des Systèmes d'Information Partagés de Santé (ASIP Santé) va dans ce sens. Elle a pour but de consolider la maitrise d'ouvrage publique afin de favoriser le développement des systèmes d'information pour la santé.

1.4 Questionnements

Le développement de tels systèmes ne va pas sans poser un certain nombre de questions éthiques et déontologiques et introduit de nouveaux dangers.

La première question est évidemment d'ordre déontologique. Le rapport Simon et Acker [Simon and Acker, 2008] résume ainsi les positions des représentations professionnelles nationale, européenne et mondiale vis à vis de la télémédecine :

- « La relation par télémédecine entre un patient et un médecin, même dans l'exercice collectif de la médecine, doit être personnalisée, c'est-à-dire reposer sur une connaissance suffisante du patient et de ses antécédents. Son consentement à ce nouveau mode d'exercice doit être obtenu. »
- « Le secret professionnel doit être garanti, ce qui oblige à un dispositif d'échange et de transmission qui soit parfaitement sécurisé. »
- « L'exercice de la télémédecine doit répondre à un besoin dont les raisons essentielles sont l'égalité d'accès aux soins, l'amélioration de la qualité des soins et de leur sécurité, objectifs

auxquels toute personne a droit, la télémédecine ayant l'avantage de raccourcir le temps d'accès et ainsi d'améliorer les chances d'un patient lorsqu'il est éloigné d'une structure de soins. »
- « La télémédecine doit se réaliser avec un dispositif technologique fiable dont les médecins sont en partie responsables. Il faut refuser de pratiquer la télémédecine si la technologie incertaine peut augmenter le risque d'erreur médicale. »

Si la circulation et le recoupement de l'information sont les fondements de ces nouvelles approches, ils soulèvent, en effet, le problème du respect de la confidentialité alors que le nombre d'intervenants s'accroit. Les progrès en matière de sécurité informatique sont un élément de réponse qui permet de faire face à d'éventuelles activités malveillantes sur les réseaux de communication. Mais le développement de la médecine électronique amène à élargir le champ des professions impliquées dans le système de soins et à confier des données médicales à des personnes qui n'ont pas nécessairement été formées au respect du secret professionnel. Si l'on définit l'équipe médicale comme l'ensemble des personnes auxquelles le patient confie son dossier celle-ci comporte désormais des personnes extérieures aux professions de santé. Se pose également la question de l'information donnée au patient. Il faut lui permettre de connaitre l'implication réelle du choix, qui doit être le sien, d'être pris en charge par un système de télémédecine.

Il réside un danger dans la complexification du système. Si l'introduction de systèmes de télémédecine peut permettre de diminuer certains risques, il est certain que ces même systèmes introduisent de nouveaux risques d'erreurs. Elles peuvent être des erreurs de saisies et, de manière plus générale, des erreurs liées à l'utilisation du système ou encore des problèmes de conception du système informatique lui-même. Il est demandé aux médecins de ne pas pratiquer la télémédecine en utilisant une technologie incertaine. Cela pose la question de l'évaluation ou de la certification de ces systèmes.

1.5 Un enjeu scientifique

Mettre à disposition de l'équipe médicale une information pertinente implique nécessairement de développer une capacité pour les systèmes à traiter l'information. La multiplication des sources de données et le partage de ces données se traduit par la création d'un volume grandissant d'informations qu'il est nécessaire de pouvoir filtrer afin que le praticien n'en soit pas submergé. L'analyse automatique de l'information apparait ainsi comme une condition nécessaire à un développement pertinent de la télémédecine. Comme nous le reverrons au travers des différentes applications présentées dans ce mémoire, l'interprétation médicale des données est une tâche difficile à automatiser pour les raisons suivantes :

- La pertinence de l'information vient souvent de la fusion de plusieurs sources de données. Les systèmes d'alertes et d'aide à la décision doivent, à l'image du raisonnement médical, se baser sur l'analyse de plusieurs facteurs. Les données à fusionner peuvent être de nature différente avec, à la fois, des paramètres qualitatifs, tels que l'appréciation de l'état général d'un patient, la présence ou l'absence d'un symptôme, et quantitatifs comme dans le cas d'un dosage biologique ou d'une mesure de poids.
- Les informations peuvent présenter des niveaux de certitude variables. La spécificité et la sensibilité d'un test permettent ainsi de mesurer la capacité à affirmer ou infirmer la présence d'une pathologie.
- Le raisonnement médical est essentiellement symbolique. Ainsi l'interprétation des mesures de pressions artérielles, qui sont des valeurs quantitatives, donne lieu à une information symbolique comme la présence ou non d'un syndrome d'hypertension artérielle.

- Les problèmes posés sont essentiellement partiellement observables et partiellement observés. Partiellement observables car, dans bien des cas, il n'existe pas de moyens techniques permettant de mesurer le paramètre que l'on cherche à estimer. Ainsi, l'état de santé d'un patient ne peut en général pas être établi de manière sûre. La seule référence est celle qui peut être donnée par le médecin mais celle-ci ne résulte pas toujours d'un consensus et peut donc varier fortement d'un médecin à l'autre. Partiellement observés car il existe souvent des données manquantes ou des facteurs extérieurs non pris en compte par le système de télémédecine.
- L'extrême complexité des processus biologiques en jeu et la vision nécessairement partielle qui peut en être donnée induit une grande variabilité des phénomènes observés. Reprenons l'exemple de la pression artérielle. Cette mesure est la résultante de plusieurs facteurs avec en particulier un facteur cardiaque (la pompe), un facteur vasculaire (le diamètre des artères) et le volume sanguin. Chaque facteur dépend lui-même de plusieurs paramètres. Pour le facteur cardiaque la fréquence cardiaque entre en jeu qui elle même dépend du système nerveux. Le volume sanguin dépend du niveau d'hydratation qui lui même dépend, entre autres, de facteurs extérieurs tels que le régime alimentaire ou la température ambiante. Le système rénine-angiotensine est un mécanisme hormonal jouant un rôle important dans la régulation de la pression artérielle agissant à la fois sur le volume sanguin et sur le facteur vasculaire. C'est un système complexe faisant intervenir notamment le rein, le foie et le poumon. Ainsi les variations observées de la pression artérielle peuvent avoir de nombreuses causes différentes et pour un facteur donné, par exemple à température ambiante constante, la mesure de pression artérielle peut présenter une grande variabilité.

2 Intelligence artificielle et médecine

L'intelligence artificielle est un domaine de recherche en informatique qui consiste à s'attaquer avec une machine de Turing à des problèmes que l'intelligence humaine sait bien résoudre, comme l'analyse de la parole, l'interprétation des images, le développement de stratégies dans les jeux.

Selon Alan Turing, une façon de préciser la question de l'intelligence d'une machine est de s'intéresser au jeu de l'imitation qui consiste, pour un interrogateur humain, à poser en aveugle des questions à une machine et à un humain pour essayer de découvrir lequel de ses deux interlocuteurs est la machine. Le jeu est, de manière générale, un terrain de compétition entre l'homme et la machine comme l'atteste l'exemple du jeu d'échecs avec les célèbres parties entre Deep Blue et Garry Kasparov.

Nous pourrions définir comme intelligent un système d'aide à la décision ou de génération d'alertes si sa validation requiert la participation d'une référence humaine. Un système d'alertes dont il serait possible de donner une spécification mathématique *a priori* et dont le but serait uniquement de vérifier cette spécification ne serait selon cette définition pas considéré comme intelligent. Ainsi un système de suivi de la pression artérielle visant à générer une alerte en cas d'hypertension, l'hypertension étant définie médicalement par un seuil numérique, ne rentre pas dans le cadre de l'intelligence artificielle. En revanche un système de suivi de la pression artérielle visant à alerter sur une aggravation des chiffres tensionnels avec une capacité à prendre en compte la variabilité entre individus et à filtrer des mesures localement élevées mais ne résultant pas d'une tendance générale pourrait être considérée comme intelligent.

2.1 Apprendre à mimer l'humain

L'un des buts de l'intelligence artificielle est d'être capable de reproduire le raisonnement humain au sens d'arriver aux même conclusions que l'humain devant un même problème, même si le cheminement de la pensée diffère. Ainsi le but de la reconnaissance de l'écriture manuscrite est d'associer à une image les mêmes symboles linguistiques qu'un lecteur humain. En ce sens, il s'agit de mimer l'humain. Il en va de même pour l'interprétation de données médicales dans le cadre des systèmes de télémédecine. L'un des enjeux est le transfert de connaissances humaines expertes dans un système informatique. Ce transfert peut se faire par un travail manuel de modélisation des connaissances de l'expert ou bien par l'utilisation d'algorithmes d'apprentissage basés sur l'analyse d'exemples traités par l'expert humain.

2.2 La nature de la connaissance

Si les fondements scientifiques de la médecine moderne, avec notamment la biologie moléculaire, la biologie cellulaire et la génétique, ont permis de très nombreuses avancées thérapeutiques, la médecine reste souvent décrite comme un art dans le sens où la connaissance du médecin repose à la fois sur un enseignement théorique et sur son expérience clinique propre qui, elle, est difficile à mettre en équations. D'autre part, il existe des connaissances médicales résultant d'observations pour lesquelles il n'existe pas de modèle explicatif théorique. La connaissance médicale est à la fois factuelle, c'est à dire résultant de la description d'observations, et théorique, c'est à dire résultant de la construction de modèles explicatifs des faits observés. L'important est, en général, de pouvoir déterminer des relations de cause à effet permettant de mener ou de rechercher des actions thérapeutiques.

La théorie des probabilités nous permet d'aborder, de manière macroscopique, des problèmes dont les mécanismes fins sont inconnus, trop complexes à modéliser ou non observables. Si cette vision macroscopique ne permet pas de produire des modèles explicatifs des processus, elle permet néanmoins d'élaborer des modèles prédictifs. Il n'est, par exemple, pas nécessaire de comprendre les mécanismes fins de l'action de l'aspirine pour pouvoir prédire son action. L'aspirine a, en effet, été utilisée depuis l'antiquité (sous la forme de décoctions de feuilles de saule) pour ses propriétés antipyrétiques et analgésiques, très longtemps avant la découverte en 1971 de son action biochimique. La modélisation stochastique permet de représenter ce type de connaissances que l'on pourrait qualifier de « statistiques » au sens où elles se fondent sur l'observation de séries d'évènements. Cependant, comme nous le verrons dans la suite de ce mémoire, cette approche permet avant tout de représenter et de manipuler, de manière quantifiée, une connaissance incertaine.

2.3 L'expérience MYCIN

Le système expert MYCIN a été développé à la fin des années 70 par Shortliffe et Buchanan à l'université de Stanford. Le but de ce système était d'identifier la bactérie à l'origine d'infections graves telles que des méningites ou des septicémies et de proposer un antibiotique avec un dosage adapté au patient. Les observations étaient entrées dans le système par l'intermédiaire d'une interface textuelle sous la forme de questions auxquelles l'utilisateur répondait par oui ou par non. Le moteur d'inférence évaluait ensuite la probabilité de présence des différents agents infectieux à l'aide d'une base de règles élémentaires. Les règles et les observations étaient assorties d'un facteur de certitude fondé sur la théorie des probabilités [Shortliffe and Buchanan, 1975],

permettant ainsi une représentation modulaire de connaissances incertaines à un moment où l'approche bayésienne paraissait difficile à mettre en œuvre. Ces facteurs de certitude ont été remis en question par la suite. Le formalisme graphique des réseaux bayésiens présente notamment l'avantage de clarifier les hypothèses d'indépendance faites sur les variables du problème, comme l'ont montré Heckerman et Shortliffe [Heckerman and Shortliffe, 1992]. D'un point de vu pratique, le système n'était pas utilisable en raison du temps trop long que nécessitait son utilisation. Concernant la qualité de ses recommandations, une évaluation consistant à comparer d'un côté les prescriptions proposées par MYCIN et de l'autre celles de prescripteurs humains a montré que le système informatique donnait des prescriptions plus adéquates [Yu *et al.*, 1979].

2.4 L'expérience Diatélic

Le système Diatélic a été développé à l'INRIA et au LORIA à la fin des années 90 pour prévenir les aggravations de l'état de santé de patients insuffisants rénaux traités par dialyse péritonéale à domicile. Il s'agit d'un système de télésurveillance médicale fondé sur une analyse automatique de fiches électroniques complétées quotidiennement par les patients. Après une première version à base de règles qui avait montré ses limites, le système d'analyse automatique a été revu en utilisant une approche bayésienne. L'état d'hydratation du patient et l'adéquation du traitement de dialyse sont en effet assimilés à un processus stochastique caché (voir chapitre 2) dont la valeur est estimée, chaque jour, à partir d'un ensemble relativement réduit d'indicateurs. Le système Diatélic peut être considéré comme une traduction probabiliste de la démarche du néphrologue prenant en compte l'évolution dans le temps des différents indicateurs ainsi que des contraintes dynamiques sur l'évolution de l'état de santé du patient. Ce système fut testé dans le cadre d'une expérimentation clinique médicale menée durant trois ans, de juin 1999 à août 2002, auprès des 30 patients : 15 patients suivis par Diatélic et 15 patients du groupe de contrôle. L'étude statistique a montré une amélioration significative dans le groupe Diatélic d'un certain nombre d'indicateurs, notamment une diminution de la consommation de médicaments hypotenseurs et une diminution du nombre de visites imprévues dans le centre de dialyse. Cette étude a également fait ressortir une tendance à la baisse du taux d'hospitalisations (nombre de jours d'hospitalisation diminué de moitié dans le groupe Diatélic) même si le nombre de patients n'était pas statistiquement suffisant pour affirmer de manière significative ce résultat. Ces bons résultats ont, néanmoins, motivé la création en 2002 de la société Diatélic SA pour diffuser le système.

3 Approches développées

Nous nous intéressons dans ce travail de thèse aux apports de la modélisation stochastique à cette problématique particulière de l'analyse de données médicales. Nous nous appuyons, en particulier, sur le formalisme graphique des réseaux bayésiens dynamiques, présenté en première partie de ce mémoire. Ce travail se veut une étude en largeur visitant, au travers de plusieurs applications particulières, différentes façons d'aborder le problème de l'interprétation de données médicales. Dans l'organisation de ce mémoire, nous distinguons deux types d'approche complémentaires.

Modélisation de la démarche : la première approche, qui fait l'objet de la deuxième partie de ce mémoire, consiste à modéliser explicitement les règles de décision utilisées par l'expert humain afin de reproduire sa démarche. Elle repose sur l'utilisation du formalisme graphique

à la fois comme support à la description des éléments du système étudié, ce que nous mettons en avant au travers de l'application présentée dans le chapitre 3, et comme outil permettant de réaliser l'inférence stochastique sur des problèmes de grande taille, comme nous le montrons dans le chapitre 4.

L'apprentissage automatique du modèle : la troisième partie de ce mémoire est consacrée à l'utilisation des techniques d'apprentissage pour construire des modèles à partir d'un ensemble d'exemples fournis par un expert humain. Le chapitre 5 introduit un certain nombre de notions et de techniques de la littérature auxquelles nous faisons référence dans les chapitres suivants. Les chapitres 6 et 7 présentent deux applications particulières d'apprentissage supervisé visant à classifier les observations selon les sorties attendues à partir d'une base d'exemples. Nous présentons finalement, dans le chapitre 8, une approche fondée sur la théorie du contrôle consistant à apprendre par renforcement une politique à partir de l'observation de l'expert humain dans sa pratique normale. Dans cette dernière approche la sémantique du modèle ne se trouve plus directement associée à son espace d'états, mais à un espace d'actions.

L'apport de ce travail de thèse se situe à la fois sur le plan méthodologique et sur le plan applicatif. Sur le plan méthodologique, nous proposons différentes façons d'aborder le problème de la modélisation du processus de diagnostic. Nous montrons que, selon les contraintes applicatives, différents types de réponses peuvent être données au problème de la modélisation du raisonnement médical. Sur le plan applicatif, ce travail a donné lieu à différentes réalisations logicielles fondées sur une « boite à outils » développée et complétée tout au long de cette thèse. Cependant, les résultats applicatifs de ce travail n'ont pas tous été rapportés dans l'écriture de ce mémoire.

Première partie

Un formalisme graphique

Introduction

L'adjectif stochastique est souvent présenté comme un synonyme d'aléatoire car la modélisation stochastique se fonde sur l'utilisation de variables aléatoires. Il provient du mot grec stokhastikos, lui même dérivé du mot stokhos qui désigne la notion de but. Ainsi la modélisation stochastique est une approche mathématique qui vise à faire des conjectures et à deviner l'état du système modélisé en utilisant des variables aléatoires. Elle permet une description macroscopique de phénomènes en prenant en compte une incertitude sur l'état et sur l'évolution du système et s'oppose ainsi à la modélisation déterministe.

Nous présentons dans cette première partie le cadre mathématique et théorique sur lequel nous nous appuyons dans les chapitres suivant pour modéliser le système dynamique que constitue le corps humain et mettre en place un raisonnement dans le but d'estimer son état à partir d'un ensemble d'indicateurs. Plus précisément, nous nous intéressons dans le premier chapitre, à la notion d'inférence bayésienne et au formalisme graphique des réseaux bayesiens. Le second chapitre est consacré à la modélisation de processus dynamiques sous la forme de réseaux bayésiens dynamiques (DBN) et à la question de l'inférence dans ces modèles.

Chapitre 1

Réseaux bayésiens

Sommaire

« Graphical models are a marriage between probability theory and graph theory. They provide a natural tool for dealing with two problems that occur throughout applied mathematics and engineering – uncertainty and complexity – and in particular they are playing an increasingly important role in the design and analysis of machine learning algorithms. Fundamental to the idea of a graphical model is the notion of modularity – a complex system is built by combining simpler parts. Probability theory provides the glue whereby the parts are combined, ensuring that the system as a whole is consistent, and providing ways to interface models to data. The

graph theoretic side of graphical models provides both an intuitively appealing interface by which humans can model highly-interacting sets of variables as well as a data structure that lends itself naturally to the design of efficient general-purpose algorithms.

Many of the classical multivariate probabilistic systems studied in fields such as statistics, systems engineering, information theory, pattern recognition and statistical mechanics are special cases of the general graphical model formalism – examples include mixture models, factor analysis, hidden Markov models, Kalman filters and Ising models. The graphical model framework provides a way to view all of these systems as instances of a common underlying formalism. This view has many advantages – in particular, specialized techniques that have been developed in one field can be transferred between research communities and exploited more widely. Moreover, the graphical model formalism provides a natural framework for the design of new systems. » – Michael I. Jordan, 1998

Ce texte, écrit par Michael I. Jordan en préface de [Jordan, 1998], présente particulièrement bien l'intérêt du formalisme graphique pour la modélisation stochastique ainsi que sa capacité à généraliser et unifier des approches pré-existantes. Nous présentons dans ce chapitre le formalisme des réseaux bayésiens qui, comme nous le verrons, est adapté à la problématique du traitement de données médicales en raison de sa capacité à décrire localement les relations entre variables, à faire ressortir clairement les hypothèses d'indépendance faites dans le modèle et à quantifier l'incertitude associée aux règles d'analyse.

1.1 De Bayes à la représentation graphique

1.1.1 Principales définitions

Une variable aléatoire X se définit comme une fonction $X : \omega \to X(\omega)$ qui associe des valeurs $X(\omega)$ à des événements ω pour lesquels une mesure de probabilité pourra être donnée.

La loi de probabilité $P(X)$, donne la probabilité des événements associés aux valeurs de X.

– Premier axiome : La probabilité d'un événement est un réel positif inférieur à 1.

$$0 \leq P(\omega) \leq 1$$

– Second axiome : La probabilité de l'univers, c'est à dire de l'ensemble des événements possibles associés à une expérience aléatoire, vaut 1.

$$P(\Omega) = 1$$

Etant donné la variable aléatoire discrète A prenant les valeurs $a_1, \ldots, a_N$, $P(A)$ est une distribution de probabilités sur les valeurs de A :

$$P(A) = (p_{a_1}, \ldots, p_{a_N}) \quad p_{a_i} \geq 0 \quad \sum_{i=1}^{N} p_{a_i} = 1 \tag{1.1}$$

Etant données deux variables aléatoires A et B, on appelle *loi jointe* la loi de probabilité $P(A, B)$ associée aux événement joints $A \wedge B$. Par opposition, la loi de probabilité $P(A)$ d'une variable A isolée s'appelle *loi marginale*.

De la relation 1.1 découle l'opération dite de *marginalisation* :

$$P(A) = \sum_B P(A, B) \tag{1.2}$$

Nous adoptons la notation courte $P(A) = \sum_B P(A, B)$ pour désigner la somme de la loi jointe $P(A, B)$ sur l'ensemble des valeurs possibles de B : $P(A = a) = \sum_b P(A = a, B = b)$

On définit la probabilité conditionnelle d'un événement A sachant un événement B de probabilité non nulle par :

$$P(A|B) = \frac{P(A, B)}{P(B)} \tag{1.3}$$

où $P(A, B)$ est la probabilité de l'événement joint.

On en déduit la règle de Bayes :

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{1.4}$$

1.1.2 Règle de la chaîne

Etant donné un ensemble de variables aléatoire $X_1, \ldots, X_N$, il est possible de déduire de l'équation 1.3 la décomposition suivante, connue sous le nom de règle de la chaîne :

$$P(X_1, ..., X_N) = P(X_1)P(X_2|X_1) \ldots P(X_N|X_1, \ldots, X_{N-1}) \tag{1.5}$$

1.1.3 Indépendance

Par définition, on dit que deux variables aléatoires sont indépendantes si et seulement si

$$P(A, B) = P(A)P(B) \tag{1.6}$$

Des équation 1.6 et 1.3, il est possible de déduire une définition équivalente lorsque $P(B)$ est non nulle :

$$P(A|B) = P(A) \tag{1.7}$$

La notion d'indépendance telle que décrite par l'équation 1.7 pourrait se traduire de la manière suivante : la connaissance du résultat de l'expérience aléatoire associée à la variable B n'apporte aucune information quand à l'issue de l'expérience aléatoire associée à la variable A. Un exemple classique est celui de deux lancers de dés successifs. La connaissance du résultat du premier lancer ne modifie pas les probabilités sur le deuxième lancer.

Remarquons que si deux variables aléatoires sont indépendantes, alors leur covariance est nulle. En effet, si A et B sont indépendantes alors $E(AB) = E(A)E(B)$ et donc $cov(A, B) = E(AB) - E(A)E(B) = 0$. Il est cependant important de noter que la réciproque n'est pas vraie. **Le fait que deux variables soient non corrélées n'implique pas leur indépendance.**

Nous pouvons donner le contre-exemple suivant :
- soit X une variable aléatoire discrète définie comme le résultat du lancer d'un dés à 6 faces. $P(X = 1) = P(X = 2).. = P(X = 6) = \frac{1}{6}$,
- et soit Y le gain associé à chaque résultat défini par :

– $Y = 0$ si $X = 1$ ou $X = 6$,
– $Y = 1$ si $X = 2$ ou $X = 5$,
– $Y = 2$ si $X = 3$ ou $X = 4$.

$$\begin{aligned} E(XY) &= \frac{1}{6}(1 \times 0) + \frac{1}{6}(6 \times 0) + \frac{1}{6}(2 \times 1) + \frac{1}{6}(5 \times 1) + \frac{1}{6}(3 \times 2) + \frac{1}{6}(4 \times 2) = \frac{21}{6} \\ E(X) &= \frac{1}{6} + 2\frac{1}{6} + 3\frac{1}{6} + 4\frac{1}{6} + 5\frac{1}{6} + 6\frac{1}{6} = \frac{21}{6} \\ E(Y) &= 0\frac{1}{6} + 1\frac{1}{6} + 2\frac{1}{6} + 2\frac{1}{6} + 1\frac{1}{6} + 0\frac{1}{6} = 1 \end{aligned}$$

et donc

$$cov(X, Y) = E(XY) - E(X)E(Y) = 0$$

Mais X et Y ne sont clairement pas indépendantes puisque le gain Y est complètement déterminé par le résultat de l'expérience X.

1.1.4 Indépendance conditionnelle

Etant données trois variables aléatoires A,B et C, A et B sont dites indépendantes conditionnellement à C si et seulement si

$$P(A, B|C) = P(A|C)P(B|C) \tag{1.8}$$

Il en découle avec le même raisonnement que pour l'indépendance classique que lorsque $P(B|C)$ est non nulle, une définition équivalente est :

$$P(A|B, C) = P(A|C) \tag{1.9}$$

La notation $A \perp B|C$ signifie que A et B sont indépendantes conditionnellement à C. Le terme *d-separation* est utilisé pour désigner deux variables « séparées » par la connaissance d'une troisième. Ces variables sont « séparées » car lorsque C est connue, l'information ne circule pas entre A et B. C'est à dire qu'une nouvelle connaissance de la valeur de B ne modifie en rien les probabilités sur les valeurs de A.

Pour illustration, prenons l'exemple du foyer épidémique de la grippe porcine de 2009, dite A (H1N1), situé au Mexique et d'un patient présentant les symptômes de la grippe. Considérons les trois variables binaires suivantes :
– $Mexique$ (Oui / Non) : le patient a t'il fait un voyage au Mexique ?
– $A(H1N1)$ (Oui / Non) : le patient est t'il porteur du virus de la grippe porcine ?
– $Test$ (Oui / Non) : le test de présence du virus est-il positif ?

Les variables $Mexique$ et $Test$ ne sont pas indépendantes. Avoir fait un voyage au Mexique augmente la probabilité d'avoir été en contact avec le virus et donc que le test s'avère positif. Maintenant supposons que le patient soit effectivement porteur de la grippe porcine. La probabilité que le test soit positif n'est plus liée à la question du voyage au Mexique. Elle ne dépend que des caractéristiques du test.

$$P(Test|A(H1N1)) = P(Test|A(H1N1), Mexique)$$

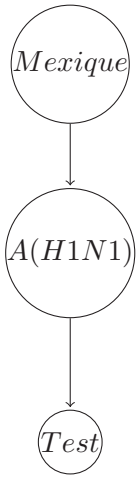


FIGURE 1.1 – Indépendance conditionnelle

1.1.5 Modèle graphique

Les réseaux bayésiens permettent de représenter graphiquement un ensemble d'indépendances conditionnelles. Un réseau bayésien est composé

- d'un ensemble de variables aléatoires $\mathcal{X} = \{X_i\}_{i=1\ldots N}$,
- du graphe dirigé acyclique $\mathcal{G} = (\mathcal{X}, Pa)$ où $Pa(X_i)$ désigne les parents de X_i dans le graphe,
- d'un ensemble de probabilités conditionnelles $\mathcal{P} = \{P(X_i|Pa(X_i))\}_{i=1\ldots N}$.

L'absence d'arcs entre noeuds du graphe encode les indépendances conditionnelles. En particulier, en numérotant les variables $(X_1, \ldots, X_N)$ par ordre hiérarchique dans le graphe nous avons $Pa(X_i) \subseteq \{X_1, \ldots, X_{i-1}\}$ pour tout $i \in [1, \ldots, N]$ et

$$P(X_i|Pa(X_i)) = P(X_i|X_1, \ldots, X_{i-1}) \tag{1.10}$$

La variable X_i est conditionnellement indépendante de $\{X_1, \ldots, X_{i-1}\} \setminus Pa(X_i)$ sachant ses parents directs $Pa(X_i)$. La règle de la chaine (1.5) devient :

$$P(X_1, \ldots, X_N) = \prod_i P\left(X_i \mid Pa(X_i)\right) \tag{1.11}$$

La loi jointe sur l'ensemble $\mathcal{X}$ des variables du réseau est représentée de manière factorisée par les paramètres $\mathcal{P}$ du réseau.

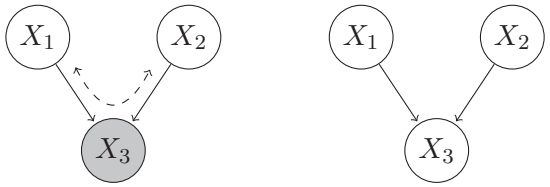
La figure 1.1 donne la représentation graphique correspondant à l'exemple précédent. Elle représente la propriété $P(Test \mid A(H1N1)) = P(Test \mid A(H1N1), Mexique)$.

1.1.6 Circulation de l'information

Nous utilisons ici le terme « circulation de l'information » pour désigner le fait qu'une nouvelle connaissance sur l'état d'une variable puisse ou non modifier la loi de probabilité sur une autre variable. Cette notion est donc directement liée aux questions d'indépendance entre les variables du réseau. L'algorithme « Bayes-Ball » [Shachter, 1998] explicite ces relations d'indépendances conditionnelles en utilisant l'image d'une balle circulant dans le réseau et en posant la définition suivante : *deux ensembles de variables A et B sont conditionnellement indépendants étant donné*

un ensemble C si et seulement si il n'existe aucun chemin pour la balle dans le réseau allant de A à B. La circulation de l'information dans le réseau dépend à la fois de la structure du réseau, mais aussi de ce qui est observé. Les règles de circulation peuvent être décrites par les trois cas de figure suivants.

1. Structure en V :



   Cette structure correspond à la factorisation suivante :

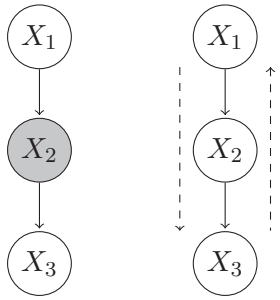
$$P(X_1, X_2, X_3) = P(X_1)P(X_2)P(X_3|X_1, X_2)$$

   Nous pouvons voir que X_1 et X_2 sont marginalement indépendantes, puisque

$$P(X_1, X_2) = \sum_{X_3} P(X_1, X_2, X_3) = P(X_1)P(X_2)$$

   mais lorsque X_3 est observée (figure de gauche), l'information circule entre X_1 et X_2. Nous pouvons donner l'illustration suivante. Considérons un jeu consistant à lancer deux dés et les trois variables suivantes : D_1 la valeur du premier dé, D_2 la valeur du second dé et $S = D_1 + D_2$ la somme des deux dés. D_1 et D_2 sont effectivement marginalement indépendantes. Mais connaissant la valeur de S, il est évident qu'une information sur D_1 nous donnera une indication sur D_2 : $P(D_2|S) \neq P(D_2|S, D_1)$.

2. Chaîne :



   Cette structure est la représentation graphique de :

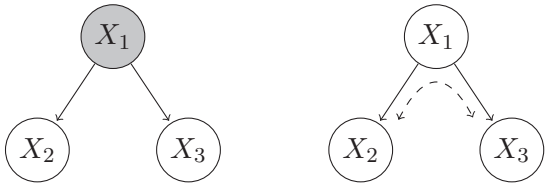
$$P(X_1, X_2, X_3) = P(X_1)P(X_2|X_1)P(X_3|X_2)$$

   Lorsque X_2 est observée l'information ne circule plus entre X_1 et X_3.

$$\begin{aligned} P(X_1, X_3|X_2) &= \frac{P(X_1)P(X_2|X_1)P(X_3|X_2)}{P(X_2)} \\ &= \frac{P(X_1, X_2)}{P(X_2)} P(X_3|X_2) \\ &= P(X_1|X_2)P(X_3|X_2) \end{aligned}$$

   X_1 et X_3 sont bien conditionnellement indépendantes relativement à X_2.

3. Un parent commun :



Cette structure est la représentation graphique de la factorisation :

$$P(X_1, X_2, X_3) = P(X_1)P(X_2|X_1)P(X_3|X_1)$$

Nous pouvons voir ici que X_2 et X_3 sont conditionnellement indépendantes relativement à X_1 puisque :

$$\begin{aligned} P(X_2, X_3|X_1) &= \frac{P(X_1)P(X_2|X_1)P(X_3|X_1)}{P(X_1)} \\ &= P(X_2|X_1)P(X_3|X_1) \end{aligned}$$

En revanche, lorsque X_1 n'est pas observée, l'information peut circuler entre X_2 et X_3. C'est le cas de deux évènements pouvant partager une même cause. Pour reprendre l'exemple du virus de la grippe, prenons les trois variables suivantes : V la présence ou non du virus, T_1 et T_2 les résultats de deux tests différents de présence du virus. Ne connaissant pas la valeur de V, le fait que T_1 soit positif vient renforcer la probabilité que T_2 soit positif aussi. Alors que connaissant la valeur de V, la probabilité sur T_2 ne dépend plus du résultat de T_1 mais simplement de la sensibilité et de la spécificité du test T_2.

1.2 Inférence bayésienne

Le terme inférence désigne de manière générale la production d'informations nouvelles à partir d'informations existantes. L'inférence en logique consiste à conclure à la vérité d'une proposition en partant de prémisses. Les prémisses sont des connaissances *a priori* sous la forme de propositions tenues pour vraies. De la même façon, l'inférence bayésienne est une démarche inductive mais travaillant cette fois sur des connaissances prenant la forme de mesures de probabilités. Elle consiste à calculer la probabilité d'une hypothèse à partir de connaissances *a priori* données sous la forme de mesures de probabilité. Dans le raisonnement bayésien, la mesure de probabilité n'est pas interprétée comme la limite à l'infini d'un calcul de fréquence mais comme la traduction numérique d'une connaissance. Elle mesure un degré de certitude dans la vérité d'une hypothèse. L'inférence bayésienne repose sur l'utilisation stricte des règles de combinaison des probabilités.

1.2.1 Un exemple simple

L'application de la règle de Bayes est le cas le plus simple d'inférence bayésienne. Prenons l'exemple d'un test médical. Notons T la variable donnant le résultat du test (positif ou négatif) et M la variable représentant l'état de la personne (malade ou non malade). Le test est caractérisé par sa sensibilité (Se) et sa spécificité (Sp). La sensibilité nous donne la probabilité que le test soit positif sachant que la personne est malade :

$$Se = P(T|M)$$

FIGURE 1.2 – Représentation graphique de la relation entre la variable M (malade ou non malade) et la variable T (test positif ou négatif) dans le cas d'un test médical.

La spécificité à l'inverse nous donne la probabilité que le test soit négatif lorsque la personne n'est pas malade :

$$Sp = P(\bar{T}|\bar{M})$$

La relation entre M et T est représentée graphiquement par la figure 1.2.

La règle de Bayes nous permet de calculer la probabilité $P(M|T)$ que la personne soit malade sachant que le test est positif.

$$P(M|T) = P(T|M)\frac{P(M)}{P(T)} \tag{1.12}$$

$$= P(T|M)\frac{P(M)}{P(T|M)P(M) + P(T|\bar{M})(1 - P(M))} \tag{1.13}$$

$$= \frac{Se\, P(M)}{Se\, P(M) + (1 - Sp)(1 - P(M))} \tag{1.14}$$

La probabilité $P(M)$ est la probabilité *a priori* que la personne soit malade. Elle est par exemple donnée par la prévalence de la maladie dans la population à laquelle appartient la personne.

Dans [Meyer *et al.*, 2009] les auteurs montrent l'importance et l'intérêt de la probabilité *a priori* dans le raisonnement bayésien au travers de l'exemple du diagnostic d'une sérologie positive pour le VIH, d'une part, chez une femme de 75 ans sans antécédents et, d'autre part, chez un toxicomane de 27 ans. Le test a les caractéristiques suivantes : $Se = 0.97$ et $Sp = 0.98$.

- Dans le cas d'une la femme âgée la prévalence de la maladie est $P(M) = 1/500000$, le calcule donne dans ce cas, $P(M|T) = 0.00097$.
- Dans le cas d'un toxicomane avec une prévalence $P(M) = 1/10$, nous obtenons $P(M|T) = 0.982$.

L'inférence bayésienne nous permet ici de fusionner l'information donnée par les caractéristiques du test avec celle donnée par la prévalence de la maladie. L'observation « test positif » est prise en compte pour calculer *a posteriori*, c'est à dire après observation du résultat du test, la probabilité que la personne soit malade. Cet exemple nous apprend que dans le cas de la femme âgée la spécificité du test n'est pas suffisante pour conclure à la présence de la maladie lorsque le test est positif.

1.2.2 *Evidences*

Dans le cadre des réseaux bayésiens, il est courant d'utiliser le terme anglais *evidence* pour désigner une observation faite sur une variable. Il faut considérer l'*evidence* comme un évènement. Dans l'exemple précédent, la variable T est une variable aléatoire binaire désignant les deux issues

possibles du test. L'*evidence* e_T est la lecture du résultat du test. C'est un évènement qui apporte une nouvelle connaissance sur la valeur de T.

Une *evidence* n'est pas nécessairement certaine. Il se peut par exemple que la lecture du résultat du test dépende d'une échelle de couleurs et qu'il y ait une incertitude sur la lecture même de ce résultat. Pearl [Pearl, 1988] utilise le terme d'*evidence* « virtuelle » pour décrire le cas où la collecte de l'*evidence* repose sur une interprétation extérieure, difficile à expliciter. Le raisonnement résultant dans la production de l'*evidence* restant caché, il est nécessaire de faire l'hypothèse d'une relation uniquement locale entre l'*evidence* et la variable sur laquelle elle porte. Celle-ci ne doit pas dépendre d'*evidences* précédentes où d'informations *a priori* déjà prises en compte dans le modèle. Dans l'exemple de l'interprétation d'une échelle de couleur pour un test médical, cela signifie par exemple que la prévalence de la maladie $P(M)$ ne doit pas être prise en compte lors de l'évaluation de l'*evidence* e_T.

L'écriture de l'*evidence* peut se faire de manière générale sous la forme de probabilités conditionnelles, ici $P(e_T|T = negatif)$ et $P(e_T|T = positif)$, quantifiant le degré de certitude attribué à chacune des valeurs de la variable observée. Selon l'hypothèse de dépendance locale nous avons $P(e_T|T, M) = P(e_T|T)$.

Remarquons que seules les valeurs relatives attribuées aux différentes hypothèses interviennent au moment de l'inférence. L'*evidence* est définie à une constante multiplicative près, au sens ou l'information apportée par $P(e_T|T)$ et $\alpha P(e_T|T)$ est la même. Sur l'exemple précédent le calcul de la probabilité *a posteriori* $P(M|e_T)$ est le suivant :

$$\begin{aligned} P(M|e_T) &= \frac{\sum_T P(e_T, T, M)}{P(e_T)} \\ &= \frac{\sum_T P(e_T|T, M)P(T|M)P(M)}{P(e_T)} \\ &= \frac{\sum_T P(e_T|T)P(T|M)P(M)}{\sum_{M,T} P(e_T|T)P(T|M)P(M)} \\ &= \frac{\sum_T \alpha P(e_T|T)P(T|M)P(M)}{\sum_{M,T} \alpha P(e_T|T)P(T|M)P(M)} \end{aligned}$$

Etant donné que $\sum_M P(M|e_T) = 1$, le dénominateur $P(e_T)$ est souvent considéré comme une constante de normalisation. Nous reverrons que la valeur $P(e_T)$ nous donne une mesure de la probabilité de l'*evidence* sachant le modèle.

L'inférence peut s'écrire dans le cas général à partir de la loi jointe en utilisant l'opération de marginalisation (1.2). Etant donné un ensemble de variables $\{X_1, \ldots, X_N\}$ et une *evidence* e portant sur la variable $X_e \in \{X_1, \ldots, X_N\}$:

$$P(X_i|e) = \frac{\sum_{X_j, j \neq i} P(e|X_e)P(X_1, \ldots, X_N)}{P(e)}$$

L'inconvénient d'utiliser la loi jointe est que celle-ci grandit exponentiellement avec le nombre de variables puisque qu'elle énumère toutes les combinaisons de valeurs possibles. Ainsi la loi jointe de N variables binaires est une table de 2^N valeurs. Le réseau bayésien nous donne une représentation factorisée (1.11), donc plus compacte, qu'il est intéressant d'exploiter.

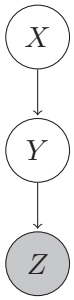


FIGURE 1.3 – Une simple chaine

1.3 Inférence dans un réseau bayésien

Il est possible d'exploiter la structure du réseau bayésien pour réaliser efficacement l'inférence dans le réseau. Prenons l'exemple d'une chaîne de trois variables X, Y et Z avec l'indépendance conditionnelle $P(Z|X,Y) = P(Z|Y)$ (figure 1.3).

Soit e_Z une *evidence* portant sur Z. Nous avons notamment les indépendances suivantes :
- $P(e_Z \mid Z,Y,X) = P(e_Z \mid Z,Y) = P(e_Z \mid Z)$ à cause de la localité de l'*evidence* ($e_Z \perp (X,Y) \mid Z$).
- $P(e_Z \mid X,Y) = P(e_Z \mid Y)$ car Y d-sépare X et Z.

Le second point peut se montrer en écrivant :

$$\begin{aligned} P(e_Z|X,Y) &= \sum_Z (P(e_Z|X,Y,Z)P(Z|X,Y)) && (1.15)\\ &= \sum_Z (P(e_Z|Y,Z)P(Z|X)) && (1.16)\\ &= P(e_Z \mid Y) && (1.17) \end{aligned}$$

En notant $\alpha = 1/P(e_Z)$ nous avons :

$$\begin{aligned} P(e_Z|Y) &= \sum_Z P(e_Z|Z)P(Z|Y) && (1.18)\\ P(e_Z|X) &= \sum_Y P(e_Z|Y)P(Y|X) && (1.19)\\ P(X|e_Z) &= \alpha P(e_Z|X)P(X) && (1.20) \end{aligned}$$

Cet exemple simple illustre le principe de la propagation de l'*evidence* le long de la structure du modèle, permettant de calculer la probabilité *a posteriori* $P(X|e_Z)$.

Ce principe de propagation peut se généraliser dans une structure d'arbre.

1.3.1 Message passing

Cette section présente brièvement l'algorithme de Pearl [Pearl, 1988] en utilisant la notation proposée par Murphy [Murphy, 1999]. L'algorithme permet de propager un ensemble d'*evidences* dans un poly-arbre (le poly-arbre est un graphe dirigé dont la version non dirigé ne présente pas de cycles). Dans un poly-arbre, un nœud peut avoir plusieurs parents. Cependant nous nous

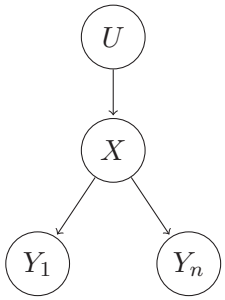


FIGURE 1.4 – Dans un graphe sans cycles chaque variable (X) sépare le graphe en un sous-graphe parent et un sous-graphe fils.

plaçons ici dans le cas plus simple où chaque nœud ne possède qu'un unique parent. Le but de cette présentation est d'introduire le principe de l'inférence basée sur une propagation locale de messages que l'on retrouve notamment dans l'algorithme forward-backward utilisé dans le cadre des HMM [Rabiner, 1989] et dans l'algorithme général d'inférence dans un réseau bayésien JLO [Lauritzen and Spiegelhalter, 1988].

Le principe de l'inférence par propagation de messages repose sur le fait que lorsque la structure du modèle ne présente pas de cycle, chaque nœud sépare l'ensemble des variables en deux sous-ensembles. Considérons la variable X avec le parent U et les fils $Y_1, \ldots, Y_N$ (voir figure 1.4). Les observations à propager portent sur un sous-ensemble E de variables ($E \subseteq \mathcal{X}$) et nous partageons l'ensemble e des *evidences* en deux sous-ensembles : l'ensemble e_X^+ des *evidences* dans le sous arbre parent, et l'ensemble e_X^- des *evidences* dans le sous-arbre fils complété de l'*evidence* portant directement sur X s'il y en a une.

Pour la variable X, nous connaissons la loi conditionnelle $P(X|U)$ (ici $Pa(X) = \{U\}$). Le sous-graphe parent apportant une information sur la variable U, Pearl utilise le terme de graphe *causal*. A l'inverse le sous-graphe fils est appelé graphe *diagnostic* car il apporte une information sur les effets de X.

L'algorithme de Pearl introduit deux messages :
- le message $\lambda_X(x) = P(e_X^-|X = x)$ remontant du graphe *diagnostic*,
- et le message $\Pi_X(x) = P(X = x|e_X^+)$ provenant du graphe *causal*

Pour calculer la loi *a posteriori* :

$$bel_X(x) = P(X = x|e) \tag{1.21}$$

L'ensemble (*sous-graphe causal*) $\rightarrow X \rightarrow$ (*sous-graphe diagnostic*) forme une chaîne et nous avons :

$$\begin{aligned} P(X, e_X^-|e_X^+) &= P(e_X^-|X, e_X^+)P(X|e_X^+) \\ &= P(e_X^-|X)P(X|e_X^+) \\ &= \lambda_X \Pi_X \\ P(X|e_X^-, e_X^+) &= P(X, e_X^-|e_X^+)/P(e_X^-|e_X^+) \end{aligned}$$

d'où

$$P(X = x|e) = \alpha \lambda_X(x)\Pi_X(x) \tag{1.22}$$

où α est la constante telle que $\sum_x P(X = x|e) = 1$.

Message λ_X

En notant $e^-_{X\to Y_1}$ les *evidences* situées en dessous de l'arc $X \to Y_1$ et e_X l'*evidence* portant sur X (si elle existe), nous pouvons calculer le message λ_X de la façon suivante en utilisant le fait que les enfants de X sont conditionnellement indépendants sachant X.

$$P(e^-_X|X=x) = P(e_X|X=x)\prod_i P(e^-_{X\to Y_i}|X=x) \tag{1.23}$$

En notant $\lambda_{X\to U}(u)$ le message $P(e^-_{U\to X}|U=u)$ que X envoie à U, l'équation (1.23) s'écrit :

$$\lambda_X(x) = \lambda_{X\to X}(x)\prod_i \lambda_{Y_i\to X}(x) \tag{1.24}$$

Une *evidence* portant sur X est prise en compte par le message $\lambda_{X\to X}(x)$ envoyé par X à lui même.

Message Π_X

Le message Π_X se calcule de la façon suivante :

$$\begin{aligned} P(X=x|e^+_X) &= \sum_u P(X=x, U=u|e^+_X) \\ &= \sum_u P(X=x|U=u)P(U=u|e^+_{U\to X}) \end{aligned}$$

En notant $\Pi_{X\to Y_i}$ le message $P(X=x|e^+_{X\to Y_i})$ que X envoie à Y_i :

$$\Pi_X(x) = \sum_u P(X=x|U=u)\Pi_{U\to X}(u) \tag{1.25}$$

Message $\lambda_{X\to U}$

Le message $\lambda_{X\to U}$ est donné par :

$$\lambda_{X\to U} = \sum_x P(X=x|U)\lambda_X(x) \tag{1.26}$$

Message $\Pi_{X\to Yi}$

Le message $\Pi_{X\to Yi}$ que X envoie à Y_i contient l'information reçue par X mais ne provenant pas de Y_i. Il est donné par :

$$\Pi_{X\to Yi}(x) = \alpha\Pi_X(x)\lambda_{X\to X}(x)\prod_{k\neq i}\lambda_{Y_k\to X(x)} \tag{1.27}$$

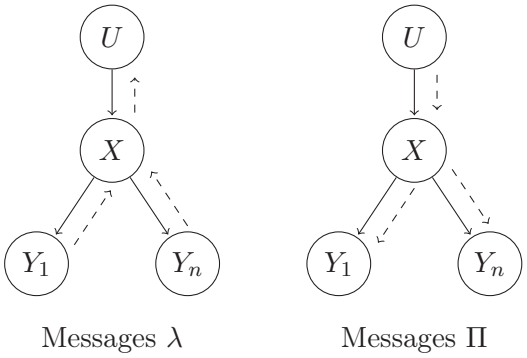


FIGURE 1.5 – Probagation des messages λ et Π dans l'arbre.

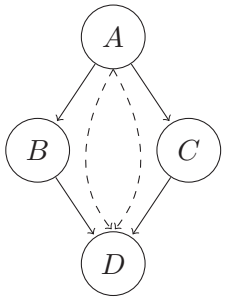


FIGURE 1.6 – L'information provenant du nœud A est propagée deux fois vers le nœud D.

Protocole de passage de message

Dans le cas particulier d'un arbre simple, l'algorithme consiste, dans une première passe, à envoyer tous les messages (λ) vers la racine puis, dans une seconde passe, à faire redescendre tous les message Pi vers les feuilles (voir figure 1.5). Lors de la seconde passe, il est possible de calculer pour chaque nœuds $bel_X(x) = P(X = x|e)$ selon l'équation (1.22).

L'algorithme de Pearl, décrit ici dans le cas d'un arbre simple, est applicable à tous les graphes dirigés dont la version non dirigée ne contient pas de cycle (ces structures sont des arbres à plusieurs têtes, ou poly-arbres). Si la variable X a plusieurs parents U_i, plusieurs messages $\lambda_{X \to U_i}$ seront construits et renvoyés vers les parents. Dans l'autre sens, les messages $\Pi_{U_i \to X}$ provenant des différents parents sont intégrés entre eux pour calculer Π_X.

L'inférence par passage de messages pose problème dans le cas où la structure présente un cycle. En effet, dans le cas de structures cycliques, nous perdons la propriété de séparation de la structure en deux parties disjointes par chaque variable et les propriétés d'indépendances conditionnelles ne sont plus vérifiées. La figure 1.6 illustre le cas où l'information provenant d'un nœud A est envoyée vers deux branches filles et se retrouve prise en compte deux fois au niveau du nœud D.

1.3.2 Arbre de jonction

L'arbre de jonction est une transformation du graphe dirigé consistant à regrouper des variables au sein de cliques de façon à faire disparaitre les cycles. Reprenons l'exemple de la structure

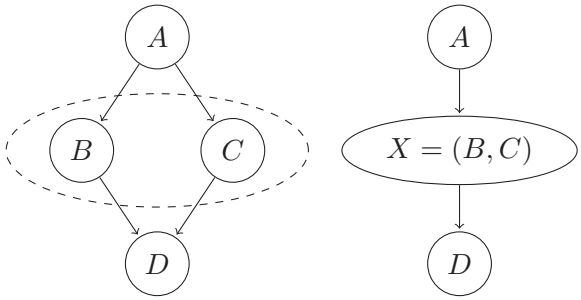


FIGURE 1.7 – Le regrouppement de B et C permet de donner une représentation équivalente sans cycle.

(A,B,C,D) représentée sur la figure 1.6. Dans cette exemple la loi jointe s'écrit :

$$P(A, B, C, D) = P(A)P(B|A)P(C|A)P(D|B, C) \tag{1.28}$$

Construisons la variable $X = (B, C)$.

$$\begin{aligned} P(X|A) &= P(B, C|A) \\ &= P(B|A)P(C|A) \end{aligned}$$

Et donc :

$$P(A, B, C, D) = P(A)P(X|A)P(D|X) \tag{1.29}$$

Ainsi, en regroupant localement les variables B et C nous pouvons donner une représentation équivalente, sans cycles, de la loi jointe comme illustré sur la figure 1.7.

L'algorithme de l'arbre de jonction, parfois appelé JLO du nom de ses auteurs Jensen, Lauritzen et Olesen [Lauritzen and Spiegelhalter, 1988] [Jensen *et al.*, 1990Pe], permet de réaliser l'inférence bayésienne dans un réseau bayésien composé de variables discrètes. Il se fonde sur la construction d'un arbre de jonction pour utiliser ensuite une propagation de messages à la manière de l'algorithme de Pearl. L'algorithme peut être étendu au cas de variables continues gaussiennes, ainsi qu'au cas, un peu plus difficile, de réseaux hybrides (composé à la fois de variables discrètes et de variables continues gaussiennes) [Lauritzen and Wermuth, 1989], [Cowell *et al.*, 1999].

Il existe plusieurs variantes de l'algorithme JLO, mais il a été montré que ces variantes d'algorithmes d'inférence exacte dans les réseaux bayésiens sont équivalents ou peuvent être dérivés d'un même algorithme [Shachter and Andersen, 1994].

Nous ne présentons ici que les grandes lignes de l'algorithme et nous recommandons la lecture de [Huang and Darwiche, 1996], qui détaille l'algorithme de l'arbre de jonction d'un point de vu procédural faisant la synthèse d'optimisations proposées dans la littérature.

Il faut distinguer dans l'algorithme de l'arbre de jonction la phase de construction de l'arbre qui résulte d'un travail sur la structure du modèle et la phase de propagation durant laquelle l'information apportée par les *evidences* circule dans l'arbre. Pour un modèle donné, l'étape de construction de l'arbre ne sera faite qu'une seule fois alors que l'étape de propagation est à répéter à chaque nouvelle inférence.

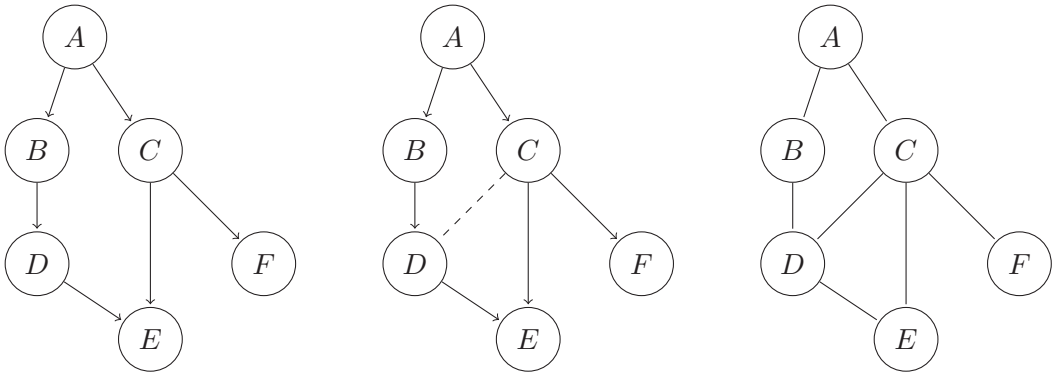


FIGURE 1.8 – Exemple de construction d'un graphe moral. Un lien non dirigé est ajouté entre les parents C et D de E. Puis la direction est retirée de tous les arcs.

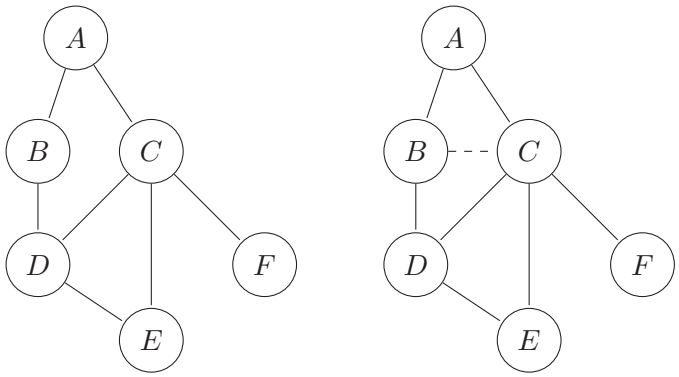


FIGURE 1.9 – Le graphe moral est triangulé en ajoutant le lien entre B et C de façon à casser le cycle de longueur 4 (A,B,C,D).

Construction de l'arbre de jonction

La construction de l'arbre de jonction commence par l'étape dite de **moralisation** du graphe. Le graphe **moral** est obtenu en « mariant », deux à deux, les parents de chaque nœud par l'ajout de liens non dirigés puis en remplaçant tous les liens dirigés du graphe par des arcs non dirigés. Un exemple de moralisation est donné sur la figure 1.8.

La seconde étape de construction de l'arbre de jonction est l'étape de **triangulation**. Le graphe triangulé est obtenu en ajoutant des arcs de manière à éliminer tous les cycles sans cordes de longueur supérieure à 3. La figure 1.9 donne un exemple de triangulation.

L'arbre de jonction est formé en regroupant les variables du graphe en un ensemble de **cliques** connectées entre elles de façon à maintenir les propriétés suivantes :

- Toute variable apparaissant dans deux cliques différentes C_i et C_j doit apparaître également dans toutes les cliques du chemin entre C_i et C_j.
- Toute famille du graphe apparaît dans au moins une clique (la famille d'une variable est ici définie comme l'ensemble composé de la variable et de ses parents dans le graphe dirigé).

L'arbre de jonction donne une nouvelle représentation factorisée de la loi jointe sous la forme d'un produit de lois locales à chaque clique. L'ensemble des variables communes à deux cliques adjacentes définit un séparateur qui matérialise l'existence d'une indépendance conditionnelle

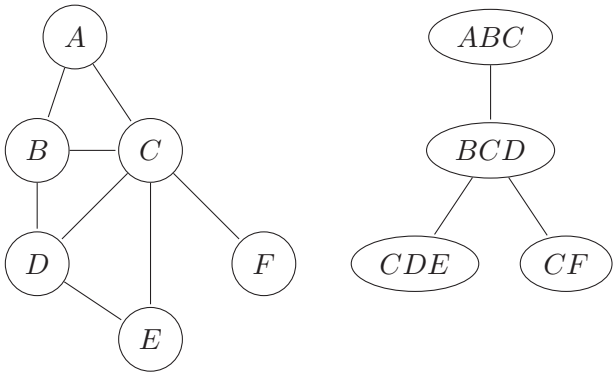


FIGURE 1.10 – Dans le graphe triangulé, le nœud F est éliminé pour former la clique (C,F). Sont ensuite formées les cliques (A,B,C), (B,C,D) et (C,D,E).

entre celles-ci. Ainsi, deux cliques de l'arbre sont conditionnellement indépendantes sachant une clique ou un séparateur intermédiaire. L'arbre de jonction est construit à partir du graphe triangulé en éliminant itérativement les nœuds et en formant des cliques composées à chaque fois du nœud éliminé et de ses voisins immédiats. Le fait de connecter entre eux les parents lors de l'étape de moralisation nous assure la constitution d'au moins une clique comportant la famille entière. La forme finale de l'arbre de jonction dépend à la fois de la triangulation qui a été faite du graphe moral et de l'ordre de constitution des cliques. La recherche de la triangulation optimale est un problème NP-complet mais il existe des heuristiques très efficaces. La figure 1.10 illustre la formation d'un arbre de jonction à partir du graphe triangulé.

Propagation des messages

Une fois l'arbre de jonction formé nous définissons pour chaque clique C_i un potentiel ϕ_{C_i}. Le potentiel est défini sur les valeurs des combinaisons des variables de la clique et il représente après propagation la loi jointe de l'ensemble des variables de la clique et des *evidences*. Nous définissons de la même façon pour chaque séparateur S_j un potentiel ϕ_{S_j}. Dans le cas de variables discrètes, les potentiels comme les probabilités conditionnelles sont représentés par une table comprenant une valeur pour chaque configuration des variables concernées.

L'ensemble des potentiels de cliques et de séparateurs permet d'écrire une nouvelle factorisation de la loi jointe sur l'ensemble $\mathcal{X}$ des variables du réseau, donnée par :

$$P(\mathcal{X}) = \frac{\prod_{C_i} \phi_{C_i}(C_i)}{\prod_{S_i} \phi_{S_i}(S_i)} \tag{1.30}$$

La propagation des messages repose sur les trois opérations élémentaires suivantes :
- L'opération de **marginalisation** qui permet de réduire le potentiel de la clique à un sous-ensemble des variables de la clique. Cette opération permettra en particulier de projeter le potentiel d'une clique vers un séparateur adjacent.
- Les opérations de **multiplication** et **division** sur les potentiels.
- L'opération d'insertion d'une ***evidence*** qui consiste à intégrer une observation dans le potentiel d'une clique.

Initialisation des potentiels Les potentiels des cliques et des séparateurs sont initialisés à 1 (l'élément neutre pour la multiplication). Chaque variable du réseau est affectée à une clique avec la contrainte que cette clique doit également contenir les parents de la variable. Rappelons que l'étape de moralisation du graphe lors de la construction de l'arbre nous assure l'existence d'une telle clique. Le potentiel de chaque clique est ensuite mis à jour en le multipliant par la loi conditionnelle des variables qui lui sont affectées.

Insertion des *evidences* Les *evidences* sont insérées pour chaque variable observée dans une seule clique en multipliant le potentiel de la clique par l'*evidence* qui, nous le rappelons, se représente par une table de valeurs sur les différents états de la variable.

Construction d'un message $C_i \to C_j$ La construction du message envoyé par la clique C_i à la clique C_j en passant par le séparateur $S_k = C_i \cap C_j$ se fait de la façon suivante :

1. Le potentiel de la clique C_i est projeté sur le séparateur en sommant le potentiel de la clique sur les variables n'appartenant pas au séparateur :

$$\phi^*_{S_k} \leftarrow \sum_{C_i \setminus S_k} \phi_{C_i} \tag{1.31}$$

2. Le potentiel de la clique C_j est multiplié par un facteur de mise à jour calculé en faisant le rapport entre l'ancien et le nouveau potentiel du séparateur :

$$\phi^*_{C_j} \leftarrow \phi_{C_j} * \frac{\phi^*_{S_k}}{\phi_{S_k}} \tag{1.32}$$

Remarquons qu'en reprenant l'interprétation du potentiel de clique comme la loi jointe sur l'ensemble des variables de la clique, cette opération correspond à $\frac{\phi_{C_j}}{\phi_{S_k}} = P(C_j|S_k)$. Ainsi le calcul du nouveau potentiel de clique ϕ_{C_j} revient à multiplier la probabilité conditionnelle $P(C_j|S_k)$ par la nouvelle probabilité $P^*(S_k)$. Il peut arriver que le potentiel du séparateur ait des valeurs nulles. Il peut être montré que $\phi_{S_k}(s_k) = 0$ seulement si $\phi^*_{S_k}(s_k) = 0$ (étant donné que les potentiels sont mis à jour par multiplication, un potentiel devenant nul ne peut que rester nul ; donc, si l'ancien potentiel est nul, le nouveau l'est aussi). Dans un tel cas, la convention habituelle est de prendre $\frac{0}{0} = 0$.

Propagation des messages A l'instar de l'algorithme de Pearl, la propagation des messages dans l'arbre de jonction peut se faire de façon séquentielle en deux phases. La première phase consiste à *collecter* les *evidences* en faisant remonter les messages depuis les feuilles. La seconde phase consiste à *distribuer* l'information depuis la racine.

Evaluation de la probabilité *a posteriori* La probabilité *a posteriori* $P(X|e)$ des états d'une variable X sachant un ensemble d'observations e se calcule simplement après le passage des messages en marginalisant le potentiel de l'une des cliques contenant X :

$$P(X|e) = \frac{\sum_{C_i \setminus X} \phi_{C_i}(C_i)}{\sum_{C_i} \phi_{C_i}(C_i)} \tag{1.33}$$

où le dénominateur $\sum_{C_i} \phi_{C_i}(C_i)$ représente la probabilité $P(e)$ de l'ensemble des *evidences*. Cette probabilité mesure la *compatibilité* entre le modèle et les *evidences*.

1.4 Utilisation de variables aléatoires continues

1.4.1 Densité de probabilité

Si X est une variable aléatoire continue définie sur $\mathbb{R}$, sa densité de probabilité est une fonction f positive ou nulle et sommable sur $\mathbb{R}$ vérifiant pour toute valeur de x :

$$P(X \leq x) = \int_{-\infty}^{x} f(u)du \tag{1.34}$$

Cette définition revient à décrire $f(x)$ comme la densité limite :

$$f(x) = \lim_{dx \to 0} \frac{P(x \leq X \leq x + dx)}{dx} \tag{1.35}$$

Il est intéressant de remarquer que puisque $P(X \in]-\infty, +\infty[) = 1$, nous avons :

$$\int_{-\infty}^{+\infty} f(u)du = 1 \tag{1.36}$$

Cependant la densité de probabilité en un point donné peut être supérieure à 1.

En dimension 2 la notion de densité peut se visualiser en considérant la limite de la probabilité pour la variable d'appartenir à une surface divisée par l'aire de cette surface lorsque celle-ci tend vers 0. La même définition peut s'exprimer en dimension 3 en considérant la probabilité d'appartenir à un volume.

1.4.2 Inférence avec des variables continues gaussiennes

Nous nous intéressons ici au cas d'un réseau composé exclusivement de variables continues gaussiennes. La loi jointe d'un ensemble $\{X_1, ...X_N\}$ de variables continues gaussiennes de dimensions $\{d_1, ..., d_N\}$ est une gaussienne multivariée de dimension $d = \sum_i d_i$.

Gaussienne multivariée

Nous désignons par variable continue gaussienne le cas d'une variable aléatoire prenant ses valeurs dans $\mathbb{R}^n$ (voir figure 1.11) dont la loi de probabilité est une loi normale $\mathcal{N}(\mu, \mathbf{\Sigma})$ de moyenne μ et de matrice de covariance $\mathbf{\Sigma}$. En dimension n la loi normale est donnée par :

$$f(\mathbf{x}; \mu, \mathbf{\Sigma}) = p \times exp\left(-\frac{1}{2}(\mathbf{x} - \mu)'\mathbf{\Sigma}^{-1}(\mathbf{x} - \mu)\right) \tag{1.37}$$

où $p = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}$ est la constante de normalisation telle que $\int_{\mathbf{x}} f(\mathbf{x}; \mu, \mathbf{\Sigma})\mathbf{dx} = 1$ ($|\Sigma|$ est le déterminant de la matrice de covariance Σ).

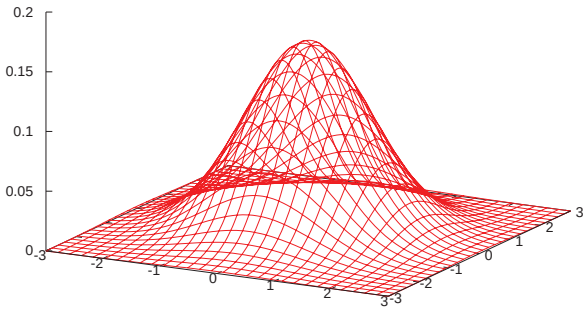


FIGURE 1.11 – Densité gaussienne en dimension 2.

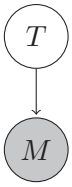


FIGURE 1.12 – Dans cette représentation graphique, la température mesurée M dépend de la vraie température T. En supposant gaussienne l'erreur de mesure, la loi P(M|T) est une gaussienne conditionnée linéairement.

Gaussienne conditionnée linéairement

Etant donné deux variables aléatoires continues gaussiennes X et Y telles que $Y = WX + \mu + Br$ ou Br est un bruit blanc gaussien de matrice de covariance Σ, la densité $P(Y|X)$ s'écrit :

$$p(Y = \mathbf{y}|X = \mathbf{x}) = f\left(\mathbf{y}; \mathbf{B}^T\mathbf{x} + \mu, \mathbf{\Sigma}\right) \tag{1.38}$$

Il s'agit d'une loi normale dont la moyenne est une fonction linéaire de la valeur x de la variable parente. Nous pouvons donner ici l'exemple de la mesure M réalisée avec un thermomètre en degrés Farenheit de la « vraie » température ambiante T qui elle est exprimée dans le système international en kelvin. En assimilant l'erreur liée à l'appareil de mesure à un bruit blanc de variance σ^2 nous pouvons écrire :

$$p(m|t) = f\left(m; \frac{9}{5}t - 459.67, \sigma^2\right) \tag{1.39}$$

La représentation graphique d'une telle relation est donnée par la figure 1.12.

Forme canonique

Il est possible de réécrire l'expression (1.37) sous une forme dite canonique en faisant remonter la constante de normalisation dans l'exponentielle. La forme canonique $\phi(\mathbf{x}; \mathbf{g}, \mathbf{h}, \mathbf{K}) = f(\mathbf{x}; \mu, \mathbf{\Sigma})$

est donnée par :

$$\phi(\mathbf{x};\mathbf{g},\mathbf{h},\mathbf{K}) = \exp\left(g + \mathbf{x}'\mathbf{h} - \frac{1}{2}\mathbf{x}'\mathbf{K}\mathbf{x}\right) \tag{1.40}$$

avec

$$\begin{aligned} \mathbf{K} &= \mathbf{\Sigma}^{-1} \\ \mathbf{h} &= \mathbf{\Sigma}^{-1}\mu \\ g &= log(p) - \frac{1}{2}\mu'\mathbf{K}\mu \end{aligned}$$

Cette forme permet d'écrire plus simplement les opérations de multiplication et de division entre distributions. Elle permet également de représenter des distributions gaussiennes conditionnées linéairement.

Représentation canonique d'une gaussienne conditionnée linéairement

La distribution

$$p(y|x) = p\exp\left[-\frac{1}{2}((y-\mu-B^Tx)^T\Sigma^{-1}(y-\mu-B^Tx))\right] \tag{1.41}$$

peut être écrite comme une fonction du vecteur $\begin{pmatrix} y \\ x \end{pmatrix}$

$$p(y|x) = \exp\left[-\frac{1}{2}\begin{pmatrix} y & x \end{pmatrix}\begin{pmatrix} \Sigma^{-1} & -\Sigma^{-1}B^T \\ -B\Sigma^{-1^T} & B\Sigma^{-1}B^T \end{pmatrix}\begin{pmatrix} y \\ x \end{pmatrix} + \begin{pmatrix} y & x \end{pmatrix}\begin{pmatrix} \Sigma^{-1}\mu \\ -B\Sigma^{-1}\mu \end{pmatrix} - \frac{1}{2}\mu^T\Sigma^{-1}\mu + \log p\right] \tag{1.42}$$

où $p = \frac{1}{(2\pi)^{n/2}\sqrt{|\Sigma|}}$ Ainsi les paramètres de la forme canonique sont :

$$g = -\frac{1}{2}\mu^T\Sigma^{-1}\mu - \frac{n}{2}\log(2\pi) - \frac{1}{2}\log|\Sigma| \tag{1.43}$$

$$h = \begin{pmatrix} \Sigma^{-1}\mu \\ -B\Sigma^{-1}\mu \end{pmatrix} \tag{1.44}$$

$$K = \begin{pmatrix} \Sigma^{-1} & -\Sigma^{-1}B^T \\ -B\Sigma^{-1^T} & B\Sigma^{-1}B^T \end{pmatrix} \tag{1.45}$$

Ce résultat présenté dans [Murphy, 2002] est une généralisation au cas vectoriel d'un résultat de [Lauritzen, 1992].

Potentiels canoniques

L'algorithme de l'arbre de jonction peut être adapté facilement au cas continu gaussien, en utilisant la forme canonique pour représenter les potentiels de cliques. Le potentiel d'une clique formée des variables $(X_1,\ldots,X_N)$ de dimensions $(w_1,\ldots,w_N)$ sera ainsi caractérisé par le triplet (g,h,K) où g est un scalaire, h un vecteur de dimension $d = \sum_i w_i$ et K une matrice $d \times d$.

Opérations de multiplication et de division Afin de pouvoir multiplier ou diviser entre elles deux fonctions de potentiels $\phi_1(x_1,\ldots,x_k;g_1,h_1,K_1)$ et $\phi_2(x_{k+1},\ldots,x_n;g_2,h_2,K_2)$, celles-ci doivent d'abord être réécrites sur un espace commun $x_1,\ldots,x_n$ en augmentant avec des zéros les vecteurs h_1,h_2 et les matrices K_1,K_2 sur les dimensions appropriées. Les opérations de multiplications et de divisions sont alors définies par :

$$(g_1,h_1,K_1)*(g_2,h_2,K_2)=(g_1+g_2,h_1+h_2,K_1+K_2) \tag{1.46}$$

et

$$(g_1,h_1,K_1)/(g_2,h_2,K_2)=(g_1-g_2,h_1-h_2,K_1-K_2) \tag{1.47}$$

Marginalisation La marginalisation consiste à calculer le potentiel ϕ_V d'un sous-ensemble de variable $V\subset W$ en sommant le potentiel ϕ_W sur les variables absentes de W ($\phi_V=\sum_{W\setminus V}\phi_W$).

$$\phi_V(y_2;g',h',K')=\int_{y_1}\phi_W\left(\begin{pmatrix}y_1\\y_2\end{pmatrix};g,\begin{pmatrix}h_1\\h_2\end{pmatrix},\begin{pmatrix}K_{11}&K_{12}\\K_{21}&K_{22}\end{pmatrix}\right) \tag{1.48}$$

où y_1 correspond aux p dimensions à sommer et y_2 aux q dimensions restantes. Le résultat de l'opération de marginalisation est alors donné par :

$$\begin{aligned} g' &= g+\frac{1}{2}\left(p\log(2\pi)-\log|K_{11}|+h_1^TK_{11}^{-1}h_1\right) && (1.49)\\ h' &= h_2-K_{21}K_{11}^{-1}h_1 && (1.50)\\ K' &= K_{22}-K_{21}K_{11}^{-1}K_{12} && (1.51)\end{aligned}$$

Insertion d'une *evidence* Plaçons nous dans le cas d'un potentiel en dimension 2 défini sur (X,Y). Ce potentiel peut se représenter comme sur la figure (1.11) par une surface. Poser l'*evidence* $Y=y$ revient à s'intéresser à l'intersection entre la surface définie par le potentiel et le plan $Y=y$. Plus généralement, le fait d'insérer une évidence dans un potentiel fait diminuer la dimension de celui-ci puisque nous nous intéressons alors uniquement à la distribution sur dimensions non observées. Le potentiel de toutes les cliques contenant la variable observée doit être réduit. Le résultat donné dans [Murphy, 2002] généralise au cas vectoriel le résultat de [Lauritzen, 1992] :

$$\phi^*(x)=\exp\left[g+\begin{pmatrix}x^T & y^T\end{pmatrix}\begin{pmatrix}h_X\\h_y\end{pmatrix}-\frac{1}{2}\begin{pmatrix}x^T & y^T\end{pmatrix}\begin{pmatrix}K_{XX}&K_{XY}\\K_{YX}&K_{YYy}\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}\right] \tag{1.52}$$

Le potentiel réduit ϕ^* est donc caractérisé par :

$$\begin{aligned} g^* &= g+h_Y^Ty-\frac{1}{2}y^TK_{YY}y && (1.53)\\ h^* &= h_X-K_{XY}y && (1.54)\\ K^* &= K_{XX} && (1.55)\end{aligned}$$

Remarquons que ce calcul correspond à l'insertion d'une *evidence* « certaine ». En reprenant l'écriture la plus générale de l'*evidence* sous la forme d'une distribution conditionnelle $p(e_Y|Y)$, une *evidence* « certaine » peut être représentée par une distribution de Dirac $\delta(y-y^*)$ centrée sur la valeur observée y^*.

Dans le cas d'une *evidence* « incertaine », la lecture de la valeur de Y peut, de la même façon, être représentée par une distribution gaussienne centrée sur la valeur lue y^* et dont la variance Σ^* caractérise alors l'incertitude. Introduire une telle évidence dans le réseau peut se faire simplement en multipliant le potentiel de l'une des cliques contenant Y par la distribution $\mathcal{N}(y^*, \Sigma^*)$ représentant l'*evidence*. Ce calcul est en fait équivalent au calcul que nous ferions en ajoutant effectivement une variable e_Y au réseau avec la loi conditionnelle $p(e_Y|y) = \mathcal{N}(e_Y - y, \Sigma^*)$ et en posant l'*evidence* certaine « $e_Y = y^*$ » comme nous l'avons vu précédemment. Cette vision présente cependant l'inconvénient d'introduire le paramètre Σ^* à l'intérieur du modèle, alors que celui-ci résulte, par définition de l'*evidence*, d'un processus externe au modèle.

Passage de messages L'initialisation des potentiels et le passage de messages se déroule comme dans le cas discret. Les potentiels sont initialisés à 1 (c'est à dire l'élément neutre pour la multiplication qui est ici le potentiel $(0, 0, 0)$). Puis la loi de chaque variable est multipliée au potentiel de la clique où celle-ci est affectée. L'insertion des *evidences* et le passage de messages sont réalisés en utilisant les opérations données précédemment.

1.5 Inférence dans un réseau hybride

Si les variables discrètes sont bien adaptées à la représentation d'un raisonnement basé sur une information symbolique (présence ou absence d'un symptôme, résultat positif ou négatif d'un test, présence ou absence d'une pathologie...) il est parfois nécessaire d'intégrer une information portant sur des variables continues. Avant de décider que le patient présente une hypertension artérielle, le médecin accède à la mesure des pressions artérielles qui est une information quantitative représentable par une variable continue. Dans le cas général le réseau bayésien peut être composé à la fois de variables discrètes et de variables continues. Nous parlons alors de réseau hybride.

1.5.1 Observations continues

Il existe un cas particulier de réseau hybride que nous rencontrons souvent lors de la modélisation d'un raisonnement humain, c'est le cas où les variables continues ne sont situées que sur des nœuds « feuilles » c'est à dire n'ayant aucun fils. Les variables continues n'ont alors que des parents discrets et la connaissance *a priori* associée à ce type de nœuds est un ensemble de loi continues $f_{Pa}(x)$ représentant les densités de probabilité de la variable X pour chacune des valeurs de ses parents Pa. Dans le cas d'une variable continue, la notation $P(X = x|Pa) = f_{Pa}(x)$ désignera un ensemble de densités de probabilité. L'introduction d'une *evidence* e_X sur un tel nœud consiste à insérer une observation du type « $X = x$ ». L'inférence revient alors à calculer :

$$P(Pa|e_X) = \alpha P(X = x|Pa)P(Pa) \tag{1.56}$$

où $\alpha = 1/P(e_X)$ est le coefficient de normalisation.

Une astuce permettant, dans ce cas particulier, d'exploiter tel quel l'algorithme d'inférence pour les réseaux discrets est d'utiliser un nœud discret X_i, à un seul état, pour chaque variable continue, et de considérer les lois $P(e_{X_i}|Pa_i)$ comme lois conditionnelles associés à ces variables. Remarquons que, comme dans le cas des *evidences* incertaines (voir section 1.2.2), seules les

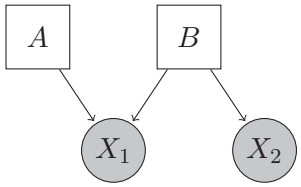


FIGURE 1.13 – Cas particulier d'un réseau où les nœuds continus X_i se trouvent sur des feuilles. Les nœuds carrés sont associés à des variables discrètes.

valeurs relatives de $f_{Pa}(x)$ pour les différentes valeurs de Pa influent sur la probabilité *a posteriori* $P(Pa|X = x)$.

1.5.2 Cas plus général

Nous considérons ici un modèle constitué de variables discrètes et continues gaussiennes avec comme contrainte qu'une variable discrète n'ait pas de parents continus. La distribution jointe d'un tel réseau hybride est un mélange de gaussiennes multivariées (à chaque configuration des variables discrètes correspond une distribution gaussienne). Les potentiels de cliques pourront donc être représentés par un ensemble de formes canoniques $(g(i), h(i), K(i))$. Nous désignons dans cette section à l'aide des lettres i, j des valeurs de variables discrètes alors que les valeurs de variables continues seront représentées par les lettres x, y.

La principale difficulté du cas hybride est que la forme gaussienne n'est pas close vis à vis de l'addition (la somme de deux gaussiennes ne peut pas se réduire à une seule gaussienne). L'addition étant utilisée au moment de l'opération de marginalisation, nous pouvons distinguer deux cas de figures concernant la marginalisation d'un potentiel hybride, la marginalisation « forte » et la marginalisation « faible ».

Marginalisation forte

Dans le cas de l'intégration de variables continues, le calcul est exact et la marginalisation est dite *forte*. C'est l'opération $\int_y \phi(i, x, y)$ décrite à la section (1.4.2) répétée pour chacune des formes canoniques du potentiel (chaque valeur de i).

Le second cas de marginalisation forte est celui de la somme de potentiels $\sum_j \phi(i, j, x)$ dont les paramètres $h(i, j)$ et $K(i, j)$ ne dépendent pas de la variable j (les distributions à sommer ont la même moyenne et la même variance). Le résultat $\phi(i, x)$ peut être représenté de manière exacte avec un nombre plus petit de formes canoniques.

$$\begin{aligned} g(i) &= \log \sum_j \exp g(i, j) \\ h(i) &= h(i, j) \\ K(i) &= K(i, j) \end{aligned}$$

Marginalisation faible

Si dans la somme $\sum_j \phi(i,j,x)$, les paramètres $h(i,j)$ et $K(i,j)$ des distributions à sommer ne sont pas indépendants des valeurs de j, le résultat est une somme de gaussiennes qu'il n'est pas possible de représenter autrement de manière exacte. C'est à dire que pour une valeur de i donnée, le potentiel $\phi(i,x) = \sum_j \phi(i,j,x)$ est une distribution multimodale. Si l'on ne s'intéresse qu'à l'espérance et à la variance de la distribution résultante, il est possible d'assimiler celle-ci à une gaussienne unique, résultat de la « fusion » des différentes gaussiennes de la somme. Cette opération s'appelle marginalisation « faible ». La marginalisation faible se fait en repassant à la forme *moment* (la forme classique) de la gaussienne $(p_{ij}, \mu_{ij}, \Sigma_{ij})$.

$$\begin{aligned} p_i &= \sum_j p_{ij} \\ \mu_i &= \sum_j \frac{p_{ij}}{p_i} \mu_{ij} \\ \Sigma_i &= \sum_j \frac{p_{ij}}{p_i} \Sigma_{ij} + (\mu_{ij} - \mu_i)(\mu_{ij} - \mu_i)^T \end{aligned}$$

Racine forte

La marginalisation forte d'un potentiel hybride est possible à condition de sommer dans l'ordre les dimensions continues puis les variables discrètes. Ceci a pour conséquence sur le passage de messages de nous obliger à organiser les nœuds de la chaine de façon à ce que la marginalisation des variables continues se face avant celle des variables discrètes lors de la collecte (remontée) des *evidences*. Cette répartition asymétrique des variables en fonction de leur nature (discrète ou continue) est décrite par le concept de racine forte [Lauritzen, 1992]. L'arbre de jonction est dit avoir une racine forte si, lorsque l'on parcourt une branche en partant de la racine, les potentiels sont « de plus en plus continus », plus précisément, si, pour deux cliques adjacentes C_1 et C_2, C_1 étant plus proche de la racine, soit C_2 ne contient pas de nouvelles variables discrètes soit le séparateur entre C_1 et C_2 est purement discret.

Lors de la phase retour (distribution des *evidences*), il est nécéssaire d'utiliser la marginalisation faible. Cependant, il peut être montré [Lauritzen, 1992] que tant que la marginalisation de la phase collecte se fait de façon forte il n'y a pas de perte d'informations au sens où après inférence :

- les potentiels sont localement cohérents car le potentiel d'un séparateur est le même quelque soit la clique adjacente à partir de laquelle celui-ci est cohérent,
- le calcule de la moyenne et de variance d'une variable réalisé par marginalisation faible sur l'une des clique est exacte,
- la loi jointe obtenue en faisant le produit des potentiels de cliques divisé par les potentiels des séparateur est exacte

$$\frac{\prod_i \phi_{C_i}}{\prod_j \phi_{S_j}}$$

.

Un arbre de jonction avec racine forte peut être obtenu simplement en imposant lors de sa construction d'éliminer les variables continues avant les variables discrètes.

1.6 Conclusion

Nous avons présenté dans ce chapitre le formalisme graphique des réseaux bayésien qui permet une description locale des relations entre un ensemble de variables aléatoires en se fondant sur la notion d'indépendance conditionnelle. Le réseau bayésien donne ainsi une représentation factorisée de la loi de probabilité jointe sur un ensemble de variable.

Le problème de l'inférence consiste à estimer les lois *a posteriori* des variables du réseau après l'insertion de nouvelles connaissances appelées *evidences*. Les *evidences* peuvent être « certaines », et correspondre à l'affirmation qu'une variable est dans un état particulier, ou « incertaines » en associant un degré de croyance à chacune des valeurs de la variable.

La structure de la représentation graphique est exploitée pour réaliser l'inférence de proche en proche, par propagation de « messages », afin de ne pas avoir recours à la loi jointe dont la taille grandit exponentiellement avec le nombre de variables. Le principe du passage de message, utilisé dans l'algorithme de Pearl dans le cas d'une structure d'arbre, a été repris dans l'algorithme JLO pour des structures quelconques en regroupant les variables en *cliques* connectées par *un arbre de jonction*. L'algorithme JLO permet ainsi de réaliser l'inférence exacte dans un réseau composé de variables discrètes ou continues gaussiennes. Dans le cas de variables continues gaussiennes, les densités de probabilité sons représentées sous une forme canonique permettant de réécrire facilement les opérations de multiplication et de division des potentiels de cliques. Le cas de réseaux hybrides est plus délicat car la somme de deux densités gaussiennes ne peux pas être réduite à une gaussienne unique. Ce cas impose notamment de réaliser les opérations de marginalisation dans un certain ordre de façon à faire disparaitre les dimensions continues avant les dimensions discrètes. Retenons cependant que, lorsque les variables continues sont situées exclusivement au niveau des feuilles, les distributions peuvent prendre une forme quelconque car ce cas est équivalent à l'utilisation d'*evidences* incertaines.

Le formalisme et les méthodes présentés dans ce chapitre ne prennent pas en compte la notion de temps. Le chapitre suivant est consacré à la modélisation de phénomènes dynamiques à l'aide de processus stochastiques. Il présente en particulier le formalisme des réseaux bayésiens dynamiques en s'appuyant sur les notions et algorithmes présentés dans ce chapitre.

Chapitre 2

Modèles de processus dynamiques

Sommaire

Il est souvent difficile d'interpréter l'état d'un patient en se basant uniquement sur une photographie de ses derniers résultats. Qu'il s'agisse d'intégrer l'évolution d'un paramètre, d'être capable de faire la différence entre une mesure localement haute et une mesure haute résultant d'une réelle évolution d'un paramètre ou bien encore de pouvoir prédire l'évolution de l'état du patient, le raisonnement médical se fonde sur une connaissance de la dynamique du système que constitue le corps humain, même si celle-ci n'est pas nécessairement formalisée mathématiquement.

Nous nous intéressons dans ce chapitre à la modélisation stochastique de processus dynamiques en utilisant le formalisme des réseaux bayésiens dynamiques (DBN) qui sont une extension des réseaux bayésiens permettant de travailler avec des processus stochastiques. Les modèles markoviens comme les chaines de markov, les modèles de markov cachés (HMM), les filtres de Kalman sont des exemples de modèles dynamiques représentables par un DBN.

Le formalisme des réseaux bayésiens dynamiques permet de donner des algorithmes génériques capables de répondre aux problèmes de l'inférence et de l'apprentissage dans ces structures.

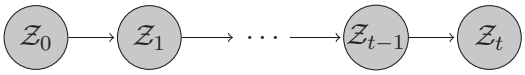


FIGURE 2.1 – Représentation graphique d'une chaîne de Markov.

2.1 Processus dynamiques markoviens

Nous nous plaçons dans le cadre de la modélisation stochastique d'un système dynamique. L'état du système est supposé décrit par un processus stochastiques $\mathcal{Z}_t$ éventuellement décomposable en un ensemble de processus

$$\mathcal{Z}_t = (Z_t^1, \ldots, Z_t^N)$$

où $t \in \mathbb{N}$ dans le cas d'un processus à temps discret.

2.1.1 Hypothèse markovienne

Un processus $\mathcal{Z}_t$ est dit markovien s'il vérifie l'indépendance conditionnelle suivante :

$$\mathcal{Z}_t \perp \mathcal{Z}_k \mid \mathcal{Z}_{t-1} \quad \text{pour tout} \quad k < t-1. \tag{2.1}$$

Cela peut s'interpréter en disant que connaître la valeur $\mathcal{Z}_k$ et $\mathcal{Z}_{t-1}$ n'apporte pas d'information supplémentaire pour prédire la valeur de $\mathcal{Z}_t$ par rapport à connaître la valeur de $\mathcal{Z}_{t-1}$ seulement. Ceci se traduit, avec le vocabulaire utilisé dans le cadre des réseaux bayésiens, par $\mathcal{Z}_{t-1}$ *d-sépare* $\mathcal{Z}_k$ et $\mathcal{Z}_t$.

Cette hypothèse d'indépendance conditionnelle, bien que correspondant souvent dans la pratique à une approximation, est fondamentale dans le sens où elle permet de décrire le système à l'aide de lois de probabilité locales et de résoudre le problème de l'inférence sans avoir besoin de recourir à la loi jointe sur l'ensemble de la séquence. En termes de lois de probabilité, l'hypothèse markovienne est caractérisée par l'équation 2.2 et représentée graphiquement par la figure 2.1

$$P(\mathcal{Z}_t \mid \mathcal{Z}_0, \ldots, \mathcal{Z}_{t-1}) = P(\mathcal{Z}_t \mid \mathcal{Z}_{t-1}) \tag{2.2}$$

Cette forme de modélisation est le pendant stochastique de la représentation donnée dans le cas déterministe d'un système dynamique d'ordre 1 par une application bijective $\phi : X \to X$ de l'espace des phases sur lui même. Pour un système dynamique déterministe à temps discret nous avons :

$$x_n = \phi(x_{n-1}) = \ldots = \phi^n(x_0) \tag{2.3}$$

Ce qui dans le cas stochastique discret s'écrit, en notant A la matrice de transition définie par $A(i,j) = P(\mathcal{Z}_t = i \mid \mathcal{Z}_{t-1} = j)$ et b_t le vecteur de probabilités donné par $b_t(i) = P(\mathcal{Z}_t = i)$:

$$P(\mathcal{Z}_n = i) = \sum_j P(\mathcal{Z}_n = i | \mathcal{Z}_{n-1} = j) P(\mathcal{Z}_{n-1} = j) \tag{2.4}$$

$$b_n = Ab_{n-1} = \ldots = A^n b_0 \tag{2.5}$$

Remarquons que cette écriture découle de la simplification de la règle de la chaine rendue possible par l'hypothèse d'indépendance conditionnelle :

$$P(\mathcal{Z}_n) = \sum_{\mathcal{Z}_0,\ldots,X_{n-1}} P(\mathcal{Z}_n \mid \mathcal{Z}_{n-1},\ldots,\mathcal{Z}_0)P(\mathcal{Z}_{n-1} \mid \mathcal{Z}_{n-2},\ldots,\mathcal{Z}_0)\ldots P(\mathcal{Z}_0) \quad (2.6)$$

$$P(\mathcal{Z}_n) = \sum_{\mathcal{Z}_0,\ldots,\mathcal{Z}_{n-1}} P(\mathcal{Z}_n \mid \mathcal{Z}_{n-1})P(\mathcal{Z}_{n-1} \mid \mathcal{Z}_{n-2})\ldots P(\mathcal{Z}_1 \mid \mathcal{Z}_0)P(\mathcal{Z}_0) \quad (2.7)$$

De la même façon qu'il est toujours possible d'augmenter la dimension de l'espace des phases d'un système dynamique pour le ramener à un système du premier ordre, il est également toujours possible d'étendre l'espace d'état d'un système pour le rendre markovien. Par exemple, si X_t décrit l'état d'un système markovien du second ordre, ayant pour loi de transition $P(X_t \mid X_{t-1}, X_{t-2})$, alors il suffit de considérer l'espace d'état $\mathcal{Z}_t = (X_t, X_{t-1})$ et la loi

$$P\left(\mathcal{Z}_t = (x_t, x_{t-1}) \mid \mathcal{Z}_{t-1} = (x'_{t-1}, x_{t-2})\right) = \delta(x_{t-1}, x'_{t-1})P(x_t \mid x_{t-1}, x_{t-2})$$

où $\delta(x_{t-1}, x'_{t-1})$ est la fonction qui vaut 1 si $x_{t-1} = x'_{t-1}$ et 0 sinon, pour avoir une représentation markovienne du même système.

2.1.2 Réseaux bayésiens dynamiques

Un réseau bayésien dynamique caractérise sous forme graphique des relations d'indépendances conditionnelles entre des processus stochastiques. En se limitant au cas markovien, le processus stochastique $\mathcal{Z}_t$ composé des variables $(Z_t^1, \ldots, Z_t^N)$ se décrit à l'aide d'une loi *a priori* $P(\mathcal{Z}_0)$ et d'une loi de transition $P(\mathcal{Z}_t \mid \mathcal{Z}_{t-1})$. Un réseau bayésien dynamique (DBN) [Murphy, 2002] est ainsi caractérisé par la paire $(B_0, B_\rightarrow)$ où B_0 est un réseau bayésien représentant la loi *a priori* $P(\mathcal{Z}_0)$ et $B_\rightarrow$ donne une représentation factorisée de la relation temporelle :

$$P(\mathcal{Z}_t \mid \mathcal{Z}_{t-1}) = \prod_i P\left(Z_t^i \mid Pa(Z_t^i)\right) \quad (2.8)$$

où $Pa(Z_t^i) \subset (\mathcal{Z}_t, \mathcal{Z}_{t-1})$ est l'ensemble des parents de Z_t^i dans le graphe.

Les DBN permettent d'unifier sous un même formalisme des modèles tels que les modèles de Markov cachés (HMM) et leurs variantes ou les filtres de Kalman. La figure 2.2 correspond à la représentation graphique d'un HMM. Il est usuel de représenter graphiquement un DBN en le « déroulant » dans le temps. Remarquons que la représentation est la même que celle d'un filtre de Kalman. La différence entre un filtre de Kalman et un HMM est la nature des variables. Dans un HMM l'état du système est donné par une variable discrète alors que celui-ci est décrit par une variable continue gaussienne dans le cas d'un filtre de Kalman. Murphy a fait une étude assez complète de la zoologie des modèles dynamiques couverts par le formalisme des DBN [Murphy, 2002].

Il est usuel de distinguer dans un modèle dynamique trois types de variables. Les variables d'état, les variables d'observations et les variables de contrôle. Cette distinction est liée en premier lieu à la position des variables dans la structure du modèle. La notion d'état est liée à la propriété markovienne du modèle. Ce sont les variables qui séparent le passé et le futur dans la représentation graphique. Les variables de contrôle sont les variables qui conditionnent la dynamique du système. Elles permettent de représenter des actions extérieures sur le système

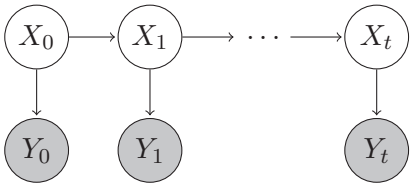


FIGURE 2.2 – Représentation graphique déroulée d'un HMM. Les variables X_t^i représentent l'état (caché) du système à chaque pas de temps. Les variables Y_t^i correspondent aux observations et dépendent directement de l'état du système.

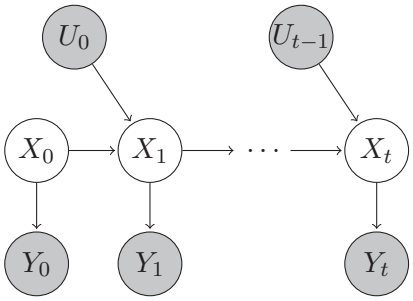


FIGURE 2.3 – Distinction entre l'état markovien X_t du système, le contrôle U_t et l'observation Y_t.

modélisé. Enfin les variables d'observation sont les variables conditionnées par l'état du système et représentant des mesures sur celui-ci. La figure 2.3 montre la représentation graphique d'un modèle dynamique partiellement observable avec une variable de contrôle. Dans un modèle partiellement observable, les variables d'état sont le plus souvent cachées même si dans le cadre des réseaux bayésiens, rien n'interdit de pouvoir observer certaines composantes de l'état du système.

Alors que dans le formalisme des HMMs ou des filtres de Kalman l'état est représenté par une variable unique, les réseaux bayésiens dynamiques permettent de considérer un ensemble de variables décrivant l'état du système. La représentation factorisée de l'espace d'état permet de réduire la complexité de problème de l'inférence comme nous l'illustrerons dans le chapitre 4. Il permet également de donner une représentation naturelle des relations entre les différents éléments constitutifs du modèle. Le formalisme graphique peut être utilisé comme support à la discussion avec l'expert humain, la structure même du modèle représentant une connaissance sur le système. Un modèle simpliste d'une infection bactérienne ou virale pourrait ainsi avoir deux variables d'état, l'une décrivant la présence ou non d'une infection virale et l'autre la présence ou non d'une infection bactérienne. La mesure de température serait alors conditionnée par ces deux variables, la fièvre étant un indicateur d'infection pour les deux causes. Dans un tel modèle, la prescription d'un antibiotique serait représentée par une variable de contrôle ne conditionnant que l'infection bactérienne comme illustrée par la figure 2.4.

Un intérêt pratique des DBN est que les algorithmes d'inférences et d'apprentissage ne sont pas spécifiques d'une structure particulières. La section suivante traite de l'inférence dans un DBN.

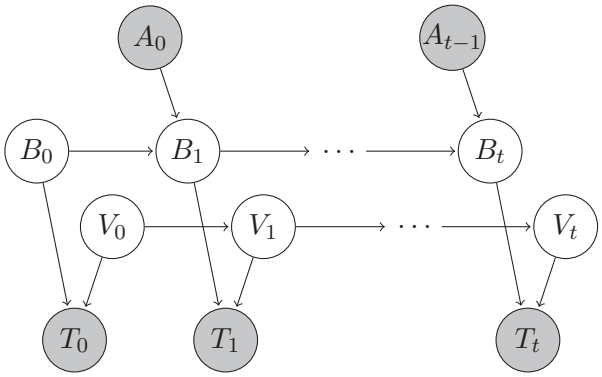


FIGURE 2.4 – Dans cet exemple, l'état du patient est représenté par les variables B_t et V_t représentant respectivement la présence ou non d'une infection bactérienne et virale. La structure du réseau décrit le fait que la prescription d'un antibiotique A_t ne conditionne que la dynamique de l'infection bactérienne.

2.2 Algorithmes d'inférence

Dans un modèle dynamique, le problème de l'inférence consiste à rechercher la distribution de probabilités sur les variables cachées du modèle à un instant donné, connaissant la séquence des observations passées et éventuellement futures. En utilisant les notations X_t et Y_t pour désigner respectivement l'ensemble des variables cachées et observées à l'instant t, le problème posé est de calculer $P(X_{t+k}|y_0, \ldots, y_t)$. Pour simplifier l'écriture, les variables de contrôle n'apparaissent pas explicitement. Dans le cadre de l'inférence dans un DBN, le contrôle peut être simplement considéré comme une observation.

Si k est positif, il s'agit de rechercher les probabilités de X dans un futur plus ou moins éloigné. Quel sera l'état du patient demain ? Les probabilités sur les observations futures peuvent également être estimées. Pour qu'un modèle permette de prédire les états futurs connaissant l'état courant, il faut que celui ci donne une connaissance suffisamment précise de la dynamique du système. Deux sources d'incertitudes [Mannor *et al.*, 2004b] interviennent rendant le problème de la prédiction difficile. La première est intrinsèque au modèle. Elle découle du non déterminisme des transitions. L'incertitude, quantifiée dans le raisonnement bayésien par l'utilisation des probabilités, tend généralement à croitre exponentiellement. Rappelons que dans une chaine de Markov de matrice de transition A, la distribution de probabilité sur les états à l'instant $t+k$ est donnée par $b_{t+k} = A^k b_t$. La seconde source d'incertitude, extrinsèque, est l'incertitude sur le modèle lui même faisant que la prédiction, juste mathématiquement par rapport au modèle peut s'avérer éloignée de la réalité.

Si $k = 0$, nous nous trouvons dans le cas du filtrage qui consiste à estimer l'état courant d'après les observations présentes et passées. Quel est l'état du patient aujourd'hui ?

Si k est négatif, la question posée est d'estimer après coup un état passé du système connaissant les observations passées et futures. Quel était l'état du patient hier ? Connaitre les observations futures permet bien évidemment d'avoir une meilleure estimation des valeurs des variables cachées. Ce problème se pose en particulier pour faire de l'apprentissage ou de l'adaptation des paramètres du modèle comme nous le verrons dans le chapitre 5. Comme nous allons le voir l'inférence dans un modèle dynamique se fait en deux phases, à l'instar du passage de messages dans le modèle statique décrit au chapitre précédent. La première phase consiste à propager vers

le futur les observations. Cette phase est suffisante pour répondre aux deux premiers problèmes ($k = 0$ et $k > 0$). Pour traiter le troisième problème $k < 0$ il est nécessaire de faire également remonter les observations récentes vers le passé.

2.2.1 *Forward-Backward*

Forward-Backward est un algorithme d'inférence bien connu dans le cadre des HMM (voir figure 2.2). Nous en rappelons ici le principe car les algorithmes d'inférence dans les DBN peuvent être vus comme des généralisations de celui-ci.

Notations

En reprenant les notations usuelles, soient :
- $\alpha_t(i) = P(X_t = i, y_0, \ldots, y_t)$ la loi jointe entre les observations passées et l'état à l'instant t,
- $\beta_t(i) = P(y_{t+1}, \ldots, y_T \mid X_t = i)$ la probabilité des observations futures sachant l'état à l'instant t,
- $\gamma_t(i) = P(X_t \mid y_0, \ldots, y_T)$ la loi *a posteriori*de l'état à l'instant t sachant l'ensemble de la séquence d'observations,
- $b_t(i) = P(y_t \mid X_t = i)$ la valeur de la fonction d'observation en y_t pour l'état $X_t = i$ (remarquons qu'ici la forme de la fonction d'observations n'a pas d'importance, nous nous trouvons dans le même cas de figure que celui présenté dans le paragraphe 1.5.1, b_t peut être considéré comme une *evidence* incertaine sur l'état X_t),
- $A(i, j) = P(X_t = i \mid X_{t-1} = j)$ la loi de transitions.

Le but de l'algorithme est de calculer les valeurs de γ_t pour $t = 0 \ldots T$. γ_t se calcule à partir de α_t et β_t :

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_j \alpha_t(j)\beta_t(j)} \tag{2.9}$$

Le dénominateur $\sum_j \alpha_t(j)\beta_t(j)$ est le coefficient de normalisation représentant la probabilité de la séquence sachant le modèle. Cette probabilité mesure la vraisemblance du modèle vis à vis de ces observations.

Phase *forward*

L'hypothèse markovienne permet de calculer les α_t itérativement. Nous avons en effet :

$$\begin{aligned}
\alpha_t(j) &= \sum_j P(X_t = j, X_{t-1} = i, y_0, \ldots, y_t) && (2.10)\\
&= \sum_j P(X_t = j, y_t \mid X_{t-1} = i, y_0, \ldots, y_{t-1})\alpha_{t-1}(i) && (2.11)\\
&= \sum_j P(X_t = j, y_t \mid X_{t-1} = i)\alpha_{t-1}(j) && (2.12)
\end{aligned}$$

car X_{t-1} d-sépare X_t, Y_t et $Y_0, \ldots, Y_{t-1}$. De la même façon, X_t d-sépare Y_t et X_{t-1} et nous avons donc :

$$\begin{aligned} P(X_t = j, y_t \mid X_{t-1} = i) &= P(y_t \mid X_t = j, X_{t-1} = i) P(X_t = j \mid X_{t-1} = i) && (2.13) \\ &= P(y_t \mid X_t = j) P(X_t = j \mid X_{t-1} = i) && (2.14) \end{aligned}$$

Il en découle :

$$\alpha_t(j) = \sum_i \alpha_{t-1}(i) A(i,j) b_t(j) \tag{2.15}$$

L'initialisation de l'algorithme s'effectue en utilisant la loi *a priori* $P(X_0)$:

$$\alpha_0(j) = b_0(i) P(X_0 = j) \tag{2.16}$$

La probabilité de la séquence sachant le modèle (vraisemblance du modèle) est donnée par :

$$P(y_0, \ldots, y_T) = \sum_j \alpha_T(j) \tag{2.17}$$

Phase *backward*

De façon similaire, il est possible de propager les observations vers le passé :

$$\begin{aligned} \beta_t(i) &= \sum_j P(X_{t+1} = j, y_{t+1}, \ldots, y_T \mid X_t = i) && (2.18) \\ &= \sum_j P(y_{t+2}, \ldots, y_T \mid X_{t+1} = j, y_{t+1}, X_t = i)) P(X_{t+1} = j, y_{t+1} \mid X_t = i) && (2.19) \\ &= \sum_j \beta_{t+1}(j) P(y_{t+1} \mid X_t = i, X_{t+1} = j) P(X_{t+1} = j \mid X_t = i) && (2.20) \end{aligned}$$

et donc

$$\beta_t(i) = \sum_j \beta_{t+1}(j) b_{t+1}(j) A(i,j) \tag{2.21}$$

Cette seconde phase s'initie en prenant $b_T(j) = 1$ pour tous les états j.

2.2.2 Algorithme de Pearl

Un HMM déroulé dans le temps ne présente pas de cycle. Nous pouvons donc réaliser l'inférence par passage de messages dans cette structure. En reprenant la notation précédente et l'équation 1.26, le message $\lambda_t = \lambda_{X_{t+1} \to X_t}$ est donné par :

$$\lambda_{X_{t+1} \to X_t}(i) = \sum_j A(i,j) \lambda_{X_{t+1}}(j) \tag{2.22}$$

où

$$\lambda_{X_{t+1}}(j) = b_{t+1}(j) \lambda_{X_{t+2} \to X_{t+1}(j)} \tag{2.23}$$

et donc

$$\lambda_t(i) = \sum_j A(i,j) b_{t+1}(j) \lambda_{t+1}(j) \tag{2.24}$$

A partir de l'équation 1.27 nous pouvons écrire le message $\Pi_t = \Pi_{X_t \to X_{t+1}}$:

$$\Pi_{X_t \to X_{t+1}}(j) = \frac{1}{p_e} b_t(j) \Pi_{X_t}(j) \tag{2.25}$$

où

$$\Pi_{X_t}(j) = \sum_i A(i,j) \Pi_{X_{t-1} \to X_t}(i) \tag{2.26}$$

et où p_e est la constante de normalisation telle que $\sum_j \Pi_{X_t \to X_{t+1}}(j) = 1$. Nous avons donc :

$$\Pi_t(j) = \frac{1}{p_e} b_t(j) \sum_i A(i,j) \Pi_{t-1}(i) \tag{2.27}$$

Ceci nous montre que l'algorithme *forward - backward* est un cas particulier de l'algorithme du passage de message de Pearl. Par définition même des messages λ et Π nous avons la correspondance suivante :

$$\lambda_{X_{t+1} \to X_t}(i) = \beta_t(i) \tag{2.28}$$

$$\Pi_{X_t \to X_{t+1}}(i) = \frac{\alpha_t(i)}{\sum_j \alpha_t(j)} \tag{2.29}$$

Il est possible de généraliser l'algorithme *forward - backward* à un DBN quelconque en utilisant l'arbre de jonction.

2.2.3 DBN déroulé

En instanciant l'ensemble des variables Z_t^i du DBN pour $t = 0, \ldots, T$, il est possible de donner une représentation déroulée du processus sous la forme d'un réseau bayésien classique. L'algorithme de l'arbre de jonction est alors utilisable. Cette option est envisageable si l'on cherche à calculer (*hors ligne*) $P(X_t \mid y_0, \ldots, y_T)$ pour $t = 0, \ldots, T$. Une fois construit l'arbre de jonction peut être réutilisé pour effectuer l'inférence avec différentes séquences d'observations $S^i = y_0^i, \ldots, y_T^i$. Les deux étapes du passage de messages (collecte et distribution des *evidences*) permettent de faire circuler les observations vers le futur et vers le passé.

Le nombre de variables étant multiplié par la longueur de la séquence T, la construction de l'arbre de jonction sera rapidement très lourde pour des grandes valeurs de T. L'algorithme de construction de l'arbre de jonction ne tenant pas compte de la structure redondante du réseau, cette méthode n'est clairement pas optimale. D'autre part, cette façon de faire n'est pas adaptée au cas du filtrage en ligne, où l'on cherche $P(X_t \mid y_0, \ldots, y_t)$ pour des valeurs continuellement plus grandes de t puisqu'il faudrait alors construire un nouveau réseau pour chaque nouvelle valeur de t.

2.2.4 Algorithme de l'interface

L'algorithme *forward - backward* exploite le fait que l'état X_t d-sépare le passé et le futur. De façon similaire, l'algorithme de l'interface [Murphy, 2002] consiste à faire ressortir un ensemble de variables I_t d-séparant le passé et le futur. L'interface I_t est composée de l'ensemble des nœuds

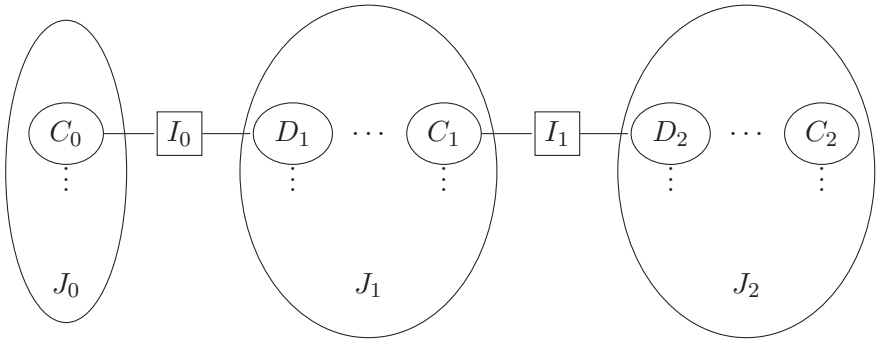


FIGURE 2.5 – Interconnexion des arbres de jonctions par l'interface I_t utilisée comme séparateur.

possédant une transition sortante vers la tranche de temps suivante. I_t d-sépare le passé du futur pour la raison suivante. Soit P un nœud du passé et F un nœud du futur, tous deux connectés à I_t. Par définition F est un nœud fils par rapport à l'interface. Il y a donc deux cas possibles, soit P est un nœud père et la structure $P \rightarrow I_t \rightarrow F$ est une chaîne, soit P est un nœud fils et la structure $P \leftarrow I_t \rightarrow F$ est une structure de type « parent commun ». Dans les deux cas, I d-sépare P et F (voir paragraphe 1.1.6).

Une fois l'interface I_t définie, l'algorithme consiste à construire une chaîne en répétant l'arbre de jonction correspondant au graphe composé des variables $\mathcal{Z}_t$ de la tranche t et de l'interface I_{t-1}. Soit G_t le graphe composé de $(I_{t-1} \cup \mathcal{Z}_t)$. L'arbre de jonction correspondant (J_t) est construit en s'assurant qu'il existe au moins une clique D_t contenant l'interface I_{t-1} et une clique C_t contenant l'interface I_t. Il suffit pour respecter cette contrainte d'ajouter un lien entre toutes les variables de I_{t-1} et celles de I_t après l'étape de moralisation durant la construction de l'arbre de jonction. Nous avons par construction $I_t = C_t \cap D_{t+1}$. La clique C_t peut alors être connectée à la clique D_{t+1} avec l'interface I_t comme séparateur comme illustrée par la figure 2.5.

La chaîne d'arbres de jonctions peut ensuite être utilisée pour réaliser l'inférence en reprenant les deux phases de l'algorithme *forward-backward*.

Phase *forward*

L'étape *forward* consiste à calculer $f_t = P(I_t|y_0, \ldots, y_t)$ à partir de $f_{t-1} = P(I_{t-1}|y_0, \ldots, y_{t-1})$ et de l'*evidence* y_t. Le message f_{t-1} est issu de la clique C_{t-1} de l'arbre J_{t-1} et est reçu par la clique D_t de l'arbre J_t. Les *evidences* y_t sont insérées dans l'abre J_t et la procédure *collecte* (voir paragraphe 1.3.2) est appelée depuis la clique C_t. f_t est calculée en marginalisant le potentiel de la clique C_t vers I_t (rappelons que $I_t \subset C_t$).

L'étape *forward* se décompose donc en différentes étapes :

1. Marginalisation du potentiel de la clique C_{t-1} pour en extraire f_{t-1}.
2. Construction de l'arbre J_t, en initialisant les potentiels avec l'élément neutre pour la multiplication (des 1 dans le cas discret).
3. Mise à jour du potentiel de D_t en le multipliant par f_{t-1}.
4. Mise à jour des potentiels à l'aide des lois conditionnelles, en multipliant la loi de chaque variable au potentiel ou celle-ci à été affectée (voir 1.3.2).
5. Multiplication des potentiels par les *evidences* y_t.

6. Collecte des *evidence* depuis la clique C_t.
7. Mémorisation des potentiels de toutes les cliques et séparateurs de J_t. Notons ϕ_t^f l'ensemble de ces potentiels.

Cette séquence d'opérations doit être répétée pour calculer successivement $f_1, f_2, \ldots, f_T$. A l'instant initial, le calcul de f_0 se fait de la même façon en collectant les *evidences* depuis la clique C_0 mais sans l'étape d'absorption du message provenant du passé. Le potentiel marginalisé doit être normalisé pour calculer f_t. Ceci permet d'éviter les problèmes liés à représentation des nombres petits. En effet, si au lieu de calculer $P(I_t|y_0, \ldots, y_t)$ nous calculions $P(I_t, y_0, \ldots, y_t)$ les valeurs seraient rapidement trop petites pour être représentables numériquement. Le coefficient de normalisation du potentiel de la clique C_t est $c_t = P(y_t|y_0, \ldots, y_t)$. Ces coefficients peuvent être conservés pour calculer la vraisemblance du modèle :

$$L_t = \log P(y_0, \ldots, y_t) = \sum_t \log(c_t) \tag{2.30}$$

Remarquons que dans cette étape nous n'avons fait que collecter les *evidences* vers C_t. Par conséquent C_t est la seule clique à avoir reçu toute l'information du passé. La distribution des *evidences* sera faite dans la phase *backward*.

Phase *backward*

Dans l'étape *backward* nous partons de ϕ_t^f et ϕ_{t+1}^b pour calculer les potentiels ϕ_t^b prenant en compte à la fois les observations passées et futures. Les *evidences* sont distribuées depuis la clique C_t, puis un message est envoyé de la clique D_t à la clique C_{t-1}. Le déroulement est le suivant :

1. Marginalisation vers I_t du potentiel de C_{t+1} (appartenant à ϕ_{t+1}^b). Ce résultat, noté b_t représente $P(I_t \mid y_0, \ldots, y_t)$.
2. Multiplication du potentiel de C_t (provenant de ϕ_t^f) par le facteur de mise à jour $\frac{f_t}{b_t}$. Autrement dit :
$$\phi_{C_t}^* = \phi_{C_t} \times \frac{\sum_{D \setminus C} \phi_{D_{t+1}}}{\sum_{C \setminus D} \phi_{C_t}} \tag{2.31}$$
3. Distribution des *evidences* depuis la clique C_t
4. Mémorisation de l'ensemble ϕ_t^b des cliques après distribution des *evidences*.

L'initialisation de la phase *backward* se fait en calculant ϕ_T^b à partir de ϕ_T^f en distribuant simplement les *evidences* depuis la clique C_T (et donc sans absorber de message).

2.2.5 Inférence approchée

Il n'est pas toujours possible de maintenir ou de raisonner sur une représentation formelle de la loi de probabilité des variables du réseau. Le filtrage particulaire est une technique d'inférence consistant à travailler à partir d'une représentation échantillonnée de la distribution de probabilité sur l'état X_t du système. Chaque particule x_t^i peut être considérée comme une hypothèse sur l'état du système. La densité pondérée des particules représentant la même hypothèse nous donne la probabilité de cette hypothèse comme décrit par l'équation 2.37.

L'utilisation d'un échantillonnage pondéré est connue sous le nom d'*importance sampling* en anglais qui se traduit par *échantillonnage préférentiel*. L'échantillonnage préférentiel consiste à

estimer une distribution $\pi(x)$ difficile à échantillonner à partir de l'échantillonnage d'une seconde distribution $q(x)$ en pondérant la valeur du tirage i par le rapport $w^i \propto \frac{\pi(x^i)}{q(x^i)}$.

Dans le cadre des modèles dynamiques la distribution $P(X_t \mid y_{1,\dots,t})$ est ciblée à partir d'une loi échantillonnée $q(X_t \mid y_{1,\dots,t})$. Les coefficients de pondérations sont donc :

$$w_t^i \approx \frac{P(X_t \mid y_{1,\dots,t})}{q(X_t \mid y_{1,\dots,t}} \tag{2.32}$$

Comme nous l'avons vu dans le cas de l'inférence exacte la distribution $P(X_{1,\dots,t} \mid y_{1,\dots,t})$ peut se calculer récursivement à partir de la loi de transition $P(X_t \mid X_{t-1})$ et de la fonction d'observation $P(Y_t|X_t)$. Les poids w_t^i se calcule de façon itérative en prenant :

$$w_t^i = \hat{w}_t^i \times w_{t-1}^i \tag{2.33}$$

avec

$$\hat{w}_t^i = \frac{P(y_t|x_t^i)P(x_t^i \mid x_{t-1}^i)}{q(x_t^i \mid x_{t-1}^i, y_t)} \tag{2.34}$$

Il est usuel, comme dans le cas de l'algorithme *Condensation*, d'échantillonner la loi dynamique du système ($q(x_t^i \mid x_{t-1}^i, y_t) = P(x_t^i|x_{t-1}^i)$). Cette technique itérative, nommée en anglais *Sequential Importance Sampling*, présente l'inconvénient que les chances de validité d'une hypothèse à long terme sont le plus souvent très faibles ce qui se traduit par une chute rapide du nombre de particules avec un poids non nul. Il est donc nécessaire de rééchantillonner régulièrement la distribution en recentrant les particules sur les hypothèses les plus probables.

Condensation

L'algorithme *Condensation* de Isard et Black [Isard and Blake, 1998] permet de faire évoluer une population de particules modélisant un processus markovien partiellement observable dans lequel l'état X_t à l'instant t est estimé à partir d'une séquence observées $y_{1,\dots,t}$. Il permet d'estimer la distribution *a posteriori* $\alpha_t(x)$:

$$\alpha_t(x) = P(x_t \mid y_{1,\dots,t}) \tag{2.35}$$

$$\alpha_t(x_t) \propto P(y_t \mid x_t) \sum_{x_{t-1}} P(x_t \mid x_{t-1})\alpha(x_{t-1}) \tag{2.36}$$

en travaillant sur une représentation échantillonnée de la loi :

$$\alpha_t(x) \approx \sum_{i=1}^{N} w_t^i \delta_{x_t^i}(x) \tag{2.37}$$

où $\delta_{x_t^i}$ est la distribution de Dirac centrée sur x_t^i. Dans l'algorithme Condensation, la distribution est rééchantillonnée à chaque pas de temps. Il se décrit en trois étapes.

1. Etape de *sélection* : la distribution α_{t-1} est rééchantillonnée avec une méthode de Monte Carlo. Chaque particule x_{t-1}^i est tirée un certain nombre de fois ou pas du tout selon son poids w_{t-1}^i.

2. L'étape de *diffusion* consiste à déplacer les particules sélectionnées en suivant la dynamique du système ($q(x_t^i \mid x_{t-1}^i, y_t) = P(x_t^i|x_{t-1}^i)$). Elle correspond à l'étape de prédiction dans les filtres de Kalman et permet d'estimer :

$$P(x_t \mid y_{1,\dots,t-1}) = \sum_{x_{t-1}} P(x_t \mid x_{t-1})\alpha(x_{t-1}) \tag{2.38}$$

   A ce stade nous avons :

$$P(x_t \mid y_{1,\dots,t-1}) \approx \frac{1}{N}\sum_{i=1}^{N} \delta_{x_t^i}(x) \tag{2.39}$$

3. Etape de *mesure* : les poids w_t^i associés aux particules sont ensuite calculés et normalisés d'après le facteur $P(y_t \mid x_t^i)$ afin de prendre en compte l'observation y_t. Cette phase est une phase de correction de la prédiction. Puisque $q(x_t^i \mid x_{t-1}^i, y_t) = P(x_t^i|x_{t-1}^i)$, nous avons d'après l'équation 2.34 :

$$w_t^i = \frac{P(y_t \mid x_t^i)}{\sum_{j=1}^{N} P(y_t \mid x_t^j)} \tag{2.40}$$

   Nous avons alors :

$$P(x_t \mid y_{1,\dots,t}) \approx \sum_{i=1}^{N} w_t^i \delta_{x_t^i}(x) \tag{2.41}$$

Selon McCormick et Isard, l'efficacité du filtre particulaire peut être décrite grâce à deux paramètres [McCormick and Isard, 2000] :

- le nombre D compris entre 1 et N de particules survivant à l'étape de sélection qui doit être suffisant pour maintenir une certaine diversité dans la population de particules,
- et le taux de survie α ($alpha \leq 1$) des particules à la sélection qui donne une indication de la qualité de l'estimation, un *alpha* petit caractérise souvent une estimation peu précise.

Le nombre de particules nécessaires à la convergence de l'algorithme grandit exponentiellement avec la dimension d de la variable d'état X_t. Pour une valeur minimum D_{min} de D et un α donnés ce nombre peut être estimé par :

$$N_{min} \geq \frac{D_{min}}{\alpha^d} \tag{2.42}$$

Généralisation

De façon plus générale, dans un DBN l'état et les observations à l'instant t sont représentés par un ensemble de variables $\mathcal{Z}_t = (Z_t^1, \dots, Z_t^M)$. Chaque variable Z_t^k peut être soit une variable observée soit une variable d'état. Nous nous situons dans le cadre d'un processus markovien du premier ordre et donc $Pa(Z_t^k) \in \{\mathcal{Z}_{t-1}, \mathcal{Z}_t\}$.

Nous présentons ici l'algorithme de filtrage particulaire (algorithme 1) tel qu'il est décrit dans [Murphy, 2002]. Il s'appuie sur la méthode de [Shachter and Peot, 1990] et [Fung and Chang, 1990] pour intégrer les *evidences* dans l'échantillonnage de la distribution. Dans cet algorithme, les étapes de diffusion et de mesure sont entremêlées. Il consiste en effet pour chaque particule i à parcourir l'ensemble des variables Z_t^k par ordre topologique. Si la variable rencontrée est une variable observée, le poids $\hat{w}_t^i$ de la particule est mis à jour. Si la variable rencontrée est une variable cachée, une valeur de la variable est tirée selon sa loi conditionnelle pour compléter le

vecteur d'état z_t^i de la particule. Un critère estimant le nombre N_{eff} de particules efficaces est utilisé pour décider de rééchantillonner ou non la distribution à chaque pas de temps :

$$N_{eff} = \frac{1}{\sum_i (w_t^i)^2} \tag{2.43}$$

Entrées : Une distribution échantillonnée $\{z_{t-1}^i, w_{t-1}^i\}_{i=1}^N$ à l'instant $t-1$ et un vecteur d'observations y_t à l'instant t
pour $i = 1$ *à* N **faire**
  $\hat{w}_t^i = 1$;
  z_t^i = vecteur vide de longueur N;
  pour *Chaque nœud k par ordre topologique* **faire**
    Soit u la valeur de $Pa(Z_t^k)$ dans (z_{t-1}^i, z_t^i) ;
    si Z_t^k *n'est pas observée* **alors**
      Compléter le vecteur z_t^i avec une valeur tirée selon la loi $P(Z_t^k \mid Pa(Z_{t-1}^k) = u)$.;
    sinon
      Compléter le vecteur z_t^i avec la valeur de Z_t^k dans y_t;
      $\hat{w}_t^i = \hat{w}_t^i \times P(Z_t^k = z_t^i \mid Pa(Z_t^k) = u)$;
    fin
  fin
  Calculer $w_t^i = \hat{w}_t^i \times w_{t-1}^i$;
fin
Calculer $w_t = \sum_i w_t^i$;
Normaliser les w_t^i : $w_t^i = \frac{w_t^i}{w_t}$;
Calculer $N_{eff} = \frac{1}{\sum_i (w_t^i)^2}$;
si $N_{eff} < seuil$ **alors**
  Sélectionner les particules selon leur poids;
  Uniformiser les poids : $w_t^i = \frac{1}{N}$;
fin

Algorithme 1: Echantillonnage de la distribution $P(\mathcal{Z}_t \mid \mathcal{Z}_{t-1}, y_t)$ et mesure du poids des particules

2.3 Conclusion

L'extension du formalisme des réseaux bayésiens au cas de processus dynamiques se fait naturellement en représentant le processus de façon « déroulée » puisque celui-ci est alors vu comme un ensemble de variables aléatoires Z_t^k indexées par le temps. La représentation « déroulée » peut être utilisée pour réaliser l'inférence lorsque le nombre de pas de temps considérés est fixe. Nous utilisons cette méthode dans le chapitre 3.

La principale difficulté posée par la modélisation de processus dynamiques est que le nombre de variables à considérer grandit avec la longueur du processus étudié. En particulier, dans le cas du filtrage en ligne, ce nombre grandit indéfiniment. L'hypothèse markovienne nous permet cependant de réduire la description du processus à une loi *a priori* B_0 et une loi de transition

$B_{\rightarrow}$. Cette représentation compacte du processus est exploitée, dans le cas particulier d'un processus caché unique, par l'algorithme *Forward-Backward* utilisé dans les HMM mais aussi par les équations de Kalman dans le cas continu. L'algorithme de l'interface, qui se fonde sur l'algorithme JLO présenté dans le chapitre 1, permet de réaliser l'inférence dans le cas plus général d'un ensemble de processus stochastiques. Cet algorithme est utilisé dans les différents travaux présentées dans la partie III de ce mémoire.

L'inférence approchée peut être réalisée dans un DBN, nous permettant de travailler avec des distributions de forme quelconque. Nous présentons dans le chapitre 4 une adaptation de la méthode particulaire présentée dans le paragraphe 2.2.5 permettant de tirer parti de la structure du problème pour diminuer le nombre de particules dans le cas de la recherche du maximum *a posteriori* (MAP).

Les algorithmes présentés dans cette première partie ont été agrégés dans une boite à outils, développée en java au cours de ce travail de thèse, et utilisée dans le développement des applications présentées dans ce mémoire.

Conclusion

Nous avons introduit dans cette première partie le formalisme graphique des réseaux bayésiens qui nous permet de modéliser un problème donné par un ensemble de variables aléatoires dont les valeurs sont associées à des hypothèses. Dans ce formalisme, un degré de certitude quant à la vérité des différentes hypothèses est donné sous la forme de mesures de probabilité associées aux différents états des variables aléatoires. L'inférence bayésienne est le mécanisme qui nous permet de raisonner à partir de connaissances *a priori* et d'un ensemble d'observations, nommées *evidences*, pour mettre à jour nos connaissances et tirer des conclusions sous la forme de probabilités *a posteriori*. Le réseau bayésien permet de décrire localement les relations entre les différentes variables du problème et de définir des *indépendances conditionnelles*. Le réseau bayésien permet ainsi une représentation factorisée du problème dont il est possible de tirer parti pour réaliser l'inférence. Nous avons présenté le principe de l'inférence par passage de messages et en particulier l'algorithme d'inférence exacte JLO qui généralise ce principe à tous types de structures composées de variables discrètes ou continues gaussiennes avec des relations linéaires.

Ce formalisme peut être étendu à la modélisation de processus dynamiques à temps discret comme le montre le chapitre 2, consacré aux DBN. Il est ainsi possible de raisonner sur l'évolution dans le temps d'un système markovien en tenant compte de son état passé pour estimer son état présent ou prédire son état futur. Les DBN généralisent et unifient un certain nombre de modèles comme les HMM ou les filtres de Kalman. L'inférence peut être réalisée, comme dans le cas statique, par passage de messages en s'appuyant sur les indépendances conditionnelles « temporelles » liées à la propriété markovienne. L'algorithme de l'interface consiste ainsi à construire une chaîne d'arbres de jonctions pour représenter le processus et à faire circuler l'information dans la chaîne une fois en avant (phase *forward*), une fois en arrière (phase *backward*). Il repose sur l'utilisation de l'algorithme JLO pour construire l'arbre. Nous avons également présenté un algorithme particulaire permettant de réaliser l'inférence de manière approchée.

Le formalisme et les algorithmes d'inférence présentés dans cette première partie sont utilisés dans toute la suite de ce mémoire. Une contribution logicielle est que tous ces algorithmes ont été développés pendant la thèse et agrégés sous la forme d'une boite à outils java.

Deuxième partie

Inférence dans l'incertain

Introduction

La seconde partie de ce mémoire traite du problème de la modélisation explicite de la démarche médicale et de l'inférence dans l'incertain en vue de découvrir l'état d'une variable cachée ou d'estimer une trajectoire dans un espace d'états.

Le formalisme des réseaux bayésiens est utilisé comme support à représentation de connaissances humaines et les algorithmes d'inférences présentés dans la partie I permettent de déduire à partir d'un ensemble d'observations, la valeur d'une variable cachée représentant l'état du patient. La modélisation de la démarche médicale se fait en définissant, avec l'aide d'un expert humain, un ensemble de variables aléatoires représentant les différents paramètres cachés ou observés du problème. Les relations entre les différentes variables sont ensuite caractérisées par des lois de probabilité conditionnelles. Nous avons suivi cette démarche dans le cadre de plusieurs applications de télésurveillance médicale couvrant les différents traitements de l'insuffisance rénale. Nous avons ainsi construit un modèle permettant de détecter précocement les rejets ou une éventuelle toxicité du traitement en transplantation rénale. L'épuration extra-rénale a été abordée dans le cadre de deux applications de télésurveillance médicale, la première, en dialyse péritonéale, la seconde, en hémodialyse qui est l'application que nous avons choisi de présenter dans le chapitre 3 de ce mémoire.

Le nombre et la nature des variables du modèle peuvent rendre le problème de l'inférence difficile comme nous le montrons dans le chapitre 4 qui traite de l'estimation des trajectoires suivie dans l'espace par les différentes articulations d'une personne, à partir de flux video. Ces trajectoires servent ensuite à mesurer des paramètres de marche comme la longueur ou la durée des pas dans le cadre d'une application visant à estimer la qualité de la marche d'une personne à son domicile afin de prévenir les chutes, et d'aider au maintien à domicile de personnes âgées. Ce problème d'inférence est particulièrement difficile en raison de l'incertitude provenant des observations vidéos et de la complexité des mouvements observés.

Chapitre 3

Un système d'aide à la décision pour l'hémodialyse

Sommaire

Nous présentons dans ce chapitre un modèle développé à partir de connaissances expertes permettant d'estimer l'adéquation du poids sec prescrit en hémodialyse. Ce travail a été effectué dans le cadre de l'action nationale de recherche et développement Dialhemo de l'INRIA en collaboration avec des néphrologues de l'ALTIR, la société Diatélic et la société Gambro.

Nous mettons en avant la méthodologie mise en œuvre qui consiste à modéliser de manière stochastique la démarche du néphrologue. Cette modélisation repose sur la compréhension des

relations d'indépendance entre les différentes variables observées et cachées du problème et leur représentation graphique sous forme de réseau bayésien dynamique. Nous recherchons ainsi à tirer parti de l'expérience du médecin afin de pouvoir traiter le problème avec relativement peu de données, celles-ci étant réservées à la validation du modèle par l'étude des résultats obtenus sur un ensemble de cas significatifs.

3.1 Problématique

Le traitement de l'insuffisance rénale chronique est un champ d'applications clé pour la télémédecine en raison de la lourdeur des traitements et du nombre important de patients concernés. C'est un des domaines ciblé, par exemple, par le rapport Simon et Acker [Simon and Acker, 2008].

La société Diatélic a été créée en 2002 pour distribuer une solution de télésurveillance médicale destinée au suivi de patients insuffisants rénaux traités par dialyse péritonéale à domicile. Cette solution est issue notamment des travaux de recherche de Jeanpierre et Charpillet [Jeanpierre, 2002].

La dialyse péritonéale, qui consiste à utiliser le péritoine comme membrane d'échange reste cependant le traitement de dialyse minoritaire et c'est tout naturellement que l'INRIA et la société Diatélic se sont intéressés à l'hémodialyse qui, comme nous allons le voir, est une technique d'épuration extra-corporelle.

3.1.1 Les principales fonctions du rein

Le rein joue un rôle essentiel dans l'élimination de l'eau et des déchets. Le compartiment liquidien de l'organisme est composé d'eau et de nombreux électrolytes dont la composition doit être constante. En éliminant ce qui est en excès et en réabsorbant en juste proportion ce qui est nécessaire, le rein assure le maintien de cet équilibre, quels que soient les aliments et les boissons absorbées. Il filtre les liquides amenés par le courant sanguin et rejette 1,5l à 2l d'urines par 24 heures. Au total, les entrées et les sorties journalières d'eau s'équilibrent.

Le rein assure également une fonction hormonale importante. Il intervient, entre autres, sur la régulation de la pression artérielle en sécrétant la rénine, et sur la production des globules rouges en sécrétant l'érythropoïétine (EPO).

3.1.2 Insuffisance rénale

L'insuffisance rénale est l'atteinte progressive ou brutale de la fonction rénale. Toutes les maladies rénales chroniques et certaines maladies rénales aiguës incurables conduisent à l'insuffisance rénale chronique dans un délai très variable : de quelques semaines à quelques dizaines d'années. L'insuffisance rénale, quelle qu'en soit la cause, est le fruit d'une réduction du nombre de néphrons, qui est l'unité structurale et fonctionnelle du rein. Le rein peut assurer ses capacités excrétrices pendant très longtemps puisqu'il n'a besoin que de 20% de ses néphrons pour fonctionner. A partir de la destruction de 80% de ses capacités, l'insuffisance rénale chronique débute. Lorsque la réduction néphronique atteint 90%, l'insuffisance rénale chronique est dite terminale et nécessite une dialyse voire une transplantation rénale. Selon le Réseau Epidémiologie et Information en Néphrologie (REIN), le nombre de personnes recevant un traitement de suppléance

par dialyse était estimé à 37 000 en France au 1er janvier 2009. Le nombre de personnes vivant avec un greffon fonctionnel était lui estimé à 31 000.

3.1.3 L'hémodialyse

La dialyse consiste à mettre en contact à travers une membrane semi-perméable, appelée dialyseur, le sang du malade et un liquide, le dialysat, dont la composition est proche de celle du plasma normal. Le dialyseur est dit semi-perméable car il ne laisse passer que les petites et moyennes molécules. A travers cette membrane les échanges se font selon un processus de diffusion tel que les substances vont du milieu le plus concentré vers le moins concentré. Ainsi l'urée [1] ou la créatinine [2] en trop forte concentration dans le sang sont éliminées dans le dyalisat, qui n'en contient pas. Il existe deux techniques de dialyse : *l'hémodialyse* qui consiste à faire circuler le sang du malade dans un rein artificiel et la *dialyse péritonéale* où la filtration se fait à l'intérieur du corps en utilisant le péritoine comme membrane d'échange.

Chaque traitement d'hémodialyse prend normalement de trois à cinq heures et il faut habituellement trois traitements par semaine. L'hémodialyse peut être effectuée dans un centre de dialyse médicalisé mais aussi dans un centre d'autodialyse ou au domicile du patient. Le rein artificiel est programmé pour faire perdre au patient une certaine quantité d'eau en l'amenant au *poids sec* qui lui a été prescrit.

Le *poids sec* est le poids du patient lorsque celui-ci est normalement hydraté. Il est déterminé par le néphrologue de manière empirique, comme illustré sur la figure 3.1, car il n'existe pas de méthode directe pour le mesurer. Une mauvaise estimation du *poids sec* par le médecin va avoir pour conséquence de déshydrater ou d'hyperhydrater le patient pouvant parfois entrainer une hospitalisation en urgence comme dans les cas d'œdème aigu du poumon.

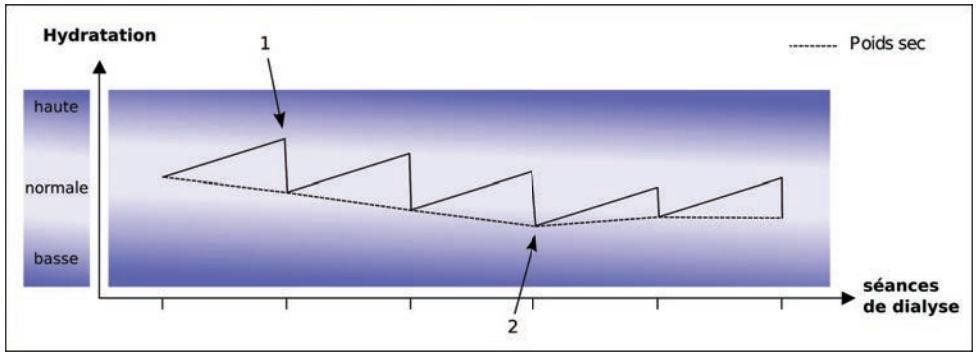


FIGURE 3.1 – Evolution de l'hydratation de séance en séance. Lorsque le patient présente des signes d'hyperhydratation au branchement (1) le médecin diminue le *poids sec*. A l'inverse, le *poids sec* doit être augmenté si le patient est déshydraté en fin de séance (2).

3.1.4 Un système expert pour le suivi du *poids sec*

Le suivi de l'adéquation du *poids sec* est un problème de régulation. L'état d'hydratation du patient peut être estimé à partir d'indicateurs comme le poids ou les pressions artérielles et,

1. Produit de dégradation du métabolisme des protéines.
2. Dérivé urinaire de la créatine phosphate qui sert de réserve énergétique aux muscles.

par ses interventions, le néphrologue corrige l'hydratation de son patient. Comme illustré sur la figure 3.2, le système d'aide à la décision a pour objectif de faciliter le retour de l'information vers le médecin afin de pouvoir proposer une surveillance continue des séances de dialyse dans les centres non médicalisés et ainsi d'y apporter la même qualité de soins que dans les centres médicalisés.

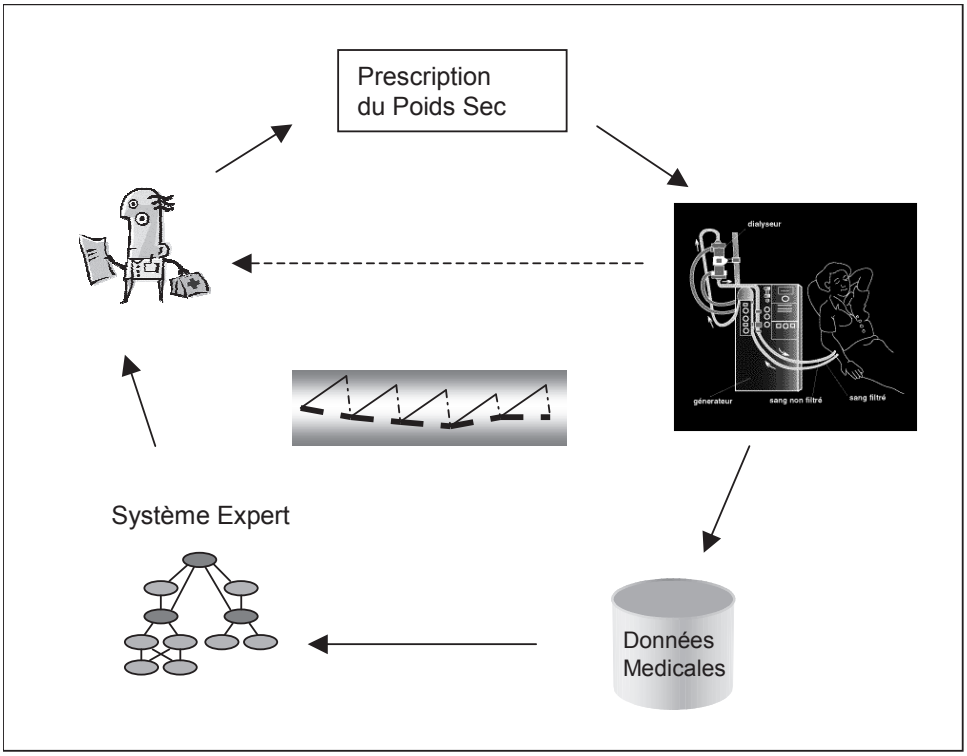


FIGURE 3.2 – Le système expert a pour but d'aider le médecin à suivre ses patients en entrant dans la boucle de régulation du *poids sec*.

La simple télétransmission de données ne constitue pas un système efficace de télésurveillance. En effet, la quantité d'information à traiter est trop importante pour que les médecins puissent consulter tous les jours les données de tous les patients. Le rôle du système expert est de sélectionner les données nécessitant un contrôle par le médecin. Cette télésurveillance ne vise pas à remplacer le contact direct entre le patient et son médecin. Elle apporte un lien complémentaire entre les visites.

3.2 Données médicales

La réalisation d'un système d'aide à la décision commence par la définition d'un ensemble d'indicateurs pertinents. L'hypothèse faite ici est qu'il est possible de produire un diagnostic sur l'adéquation du *poids sec* en se basant sur un nombre restreint d'indicateurs : le poids, la tension artérielle et les paramètres de la séance de dialyse. Un certain nombre d'autre signes cliniques, utilisés par les néphrologues pour déterminer le *poids sec*, ne sont pas pris en compte par notre système, comme par exemple la présence d'œdèmes. En effet, le système étant d'abord destiné aux centres d'autodialyse, il est fondé sur des données facilement disponibles dans ces centres.

– Le poids du patient est mesuré en début et en fin de séance d'hémodialyse. Ses variations peuvent être dues à des changements du niveau d'hydratation ou bien à des variations de la masse grasse. Le poids de fin de séance est contrôlé par le néphrologue par l'intermédiaire de la prescription du *poids sec*. Cependant il peut exister de petits écarts entre le *poids sec* prescrit et le poids de fin de séance.
– Les pressions artérielles sont mesurées en début et en fin de séance en position allongée puis en position debout. Ces pressions dépendent, entre autres, du volume sanguin qui lui même est lié à l'état d'hydratation du patient.
– L'ultrafiltration est calculée d'après les paramètres de la séance. Il s'agit de la quantité de liquide à extraire du corps du malade. Cette quantité est égale à la différence entre le poids mesuré avant la séance de dialyse et le *poids sec* estimé par le médecin. Le poids perdu par unité de temps est appelé ultrafiltration horaire :

$$UFH = \frac{Ultrafiltration}{Duree\_Seance}. \tag{3.1}$$

La quantité totale d'eau extraite pendant la séance de dialyse est appelée ultrafiltration totale (UFT). Si l'UFT ou l'UFH est trop importante, la séance risque d'être mal supportée par le patient ce qui peut se traduire, par exemple, par des hypotensions au débranchement ou des crampes.

Ces mesures sont saisies à chaque séance par le patient ou le personnel médical puis transmises par internet à l'application de télémédecine. Un exemple de tracé des différentes mesures est donné par la figure 3.3.

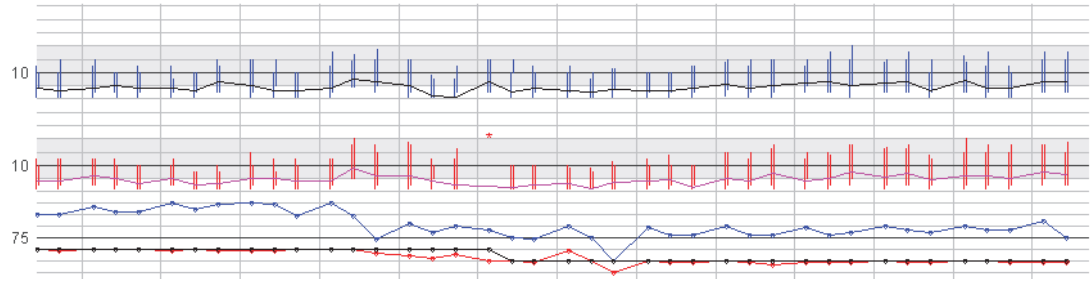


FIGURE 3.3 – Exemple de tracé des paramètres recueillis à chaque séance de dialyse. Les deux premières séries de mesures correspondent aux pressions systoliques et diastoliques mesurées en position allongée (barre de gauche) et debout (barre de droite) en début (en bleu) et en fin (en rouge) de séance. Chaque barre relie la pression systolique (point haut) à la pression diastolique (point bas). Les lignes continues noire et rose relient les pressions moyennes, en position allongée, calculées comme $\frac{2}{3}diastolique + \frac{1}{3}systolique$ respectivement en début et en fin de séance. Les deux séries suivantes sont les mesures de poids en début et en fin de séance, respectivement en bleu et en rouge. Enfin, le tracé noir représente le *poids sec* prescrit. Horizontalement, chaque case représente une semaine de traitement.

3.3 Modèle bayésien

Pour générer un diagnostic sur l'adéquation du *poids sec*, nous proposons un modèle qui « explique » les observations par rapport au processus de la dialyse. Ce modèle est le fruit d'une

collaboration étroite avec les néphrologues de l'ALTIR [3]. Il a été construit d'après les règles de diagnostic énoncées par les médecins. La solution apportée par les réseaux bayésiens consiste à décrire de manière graphique les indépendances conditionnelles qui existent entre les différentes variables observées et cachées du processus de dialyse. A partir d'un modèle *a priori* et d'une séquence d'observations nous pouvons par inférence calculer des distributions de probabilité sur les valeurs des variables cachées.

3.3.1 Les règles de diagnostic

Le diagnostic est l'identification d'une maladie par ses symptômes. Après avoir interrogé et examiné le malade, le médecin puise dans ses connaissances pour identifier la maladie en cause. Ces connaissances *a priori* peuvent se décrire sous la forme de règles, souvent incertaines.

Le suivi de la qualité du *poids sec* peut se voir comme un problème de classification où l'on cherche à déterminer les séances correspondant à un *poids sec* trop bas, normal ou trop haut. Durant les discussions avec les néphrologues, nous avons listé un certain nombre de règles utilisées pour le suivi du *poids sec* de patients insuffisants rénaux traités par hémodialyse que nous pouvons résumer comme suit.

Signes associés à un *poids sec* trop haut :
- tension élevée inhabituelle en début de séance sans prise de poids anormalement importante,
- poids de sortie (fin de séance de dialyse) inférieur au *poids sec* avec une tension normale,
- tension élevée en fin de séance.

Signes associés à un *poids sec* trop bas :
- tension basse en début de séance,
- tension basse en fin de séance,
- hypotension orthostatique [4] inhabituelle en fin de séance.

Si ces signes sont associés à un poids de sortie supérieur au *poids sec*, la probabilité que le *poids sec* soit trop bas est renforcée. D'autre part, les hypotensions et hypotensions orthostatiques en fin de séance peuvent ne pas être liées à un état de déshydratation mais être dues à une séance mal supportée à cause d'une ultrafiltration ou d'une ultrafiltration horaire trop importante.

3.3.2 Réseau bayésien pour le suivi du *poids sec*

Le réseau donné par la figure 3.4 est un graphe dirigé acyclique comprenant un ensemble de noeuds représentant les entités cliniques connectées par des arcs. Comme nous l'avons vu dans le chapitre 1, l'absence d'arc entre deux nœuds traduit l'existence d'une indépendance conditionnelle dépendant du type de structure (voir paragraphe 1.1.6). Dans le cas général, le sens de l'arc n'est pas nécessairement associé à une relation causale, la notion de causalité étant très délicate à définir et à manipuler. Cependant, dans le cas présent, nous lui donnons cette signification pour faciliter la conception et la discussion avec l'expert. Nous considérons donc

3. Association Lorraine pour le Traitement de l'Insuffisance Rénale
4. Chute de tension lors du passage de la position couchée à la position debout

que l'état d'une variable est la conséquence de l'état des variables parentes. La force de cette influence étant quantifiée par les probabilités conditionnelles associées à la variable.

Le tableau 3.1 décrit l'ensemble des variables utilisées dans le modèle représenté par la figure 3.4. La signification des différentes relations conditionnelles est précisée dans les paragraphes suivants.

Variables Cachées	**Variables Observées**
Poids sec (PS) Hydratation au branchement (Hb) Hydratation au débranchement (Hd)	Poids au branchement (Pb) Poids au débranchement (Pd) Tension au branchement (Tb) Tension au débranchement (Td) Ultrafiltration horaire (UFH) Ultrafiltration totale (UFT) Hypotension orthostatique au branchement (HOb) Hypotension orthostatique au débranchement (HOd)

TABLE 3.1 – Variables utilisées pour le suivi de la qualité du *poids sec*

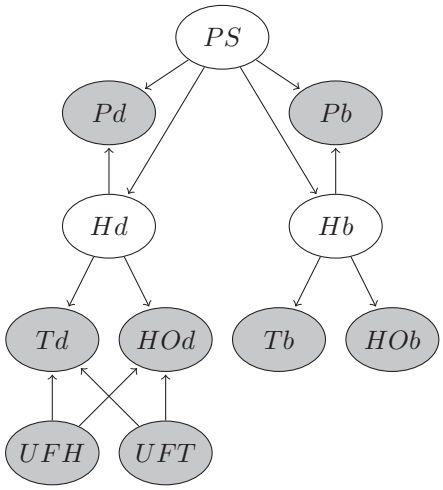


FIGURE 3.4 – Graphe causal pour le suivi du *poids sec*.

$Hb \rightarrow Tb$, $Hd \rightarrow Td$ **:** Il existe une relation de cause à effet entre l'état d'hydratation et la tension. En effet, plus le volume sanguin est grand, plus la pression sur les parois des vaisseaux sanguins est importante. L'hypertension est donc un signe d'hyperhydratation et l'hypotension est un signe de déshydratation.

$Hb \rightarrow HOb$, $Hd \rightarrow HOd$ **:** L'hypotension orthostatique est un signe de déshydratation. Lors du passage de la position couchée à la position debout, le volume sanguin a tendance à se déplacer vers le bas du corps. Normalement les vaisseaux sanguins se contractent pour maintenir la pression artérielle constante dans le haut du corps mais si le volume sanguin est trop faible, ce mécanisme ne suffit pas et la tension artérielle diminue lors du passage à la position debout. Il est à noter que certaines pathologies comme le diabète peuvent affecter ce réflexe naturel d'adaptation de la tension.

$UFH/UFT \rightarrow HOd/Td$ **:** Une ultrafiltration ou une ultrafiltration horaire trop importantes peuvent expliquer une hypotension ou une hypotension orthostatique au débranchement. En effet, tout au long de la séance l'eau est extraite du volume sanguin et il se peut, qu'en fin de séance, un déséquilibre apparaisse entre les différents compartiments liquidiens de l'organisme (intracellulaire, interstitiel et volume sanguin) en particulier si le patient avait beaucoup d'eau à perdre. L'hypotension ou l'hypotension orthostatique ne sont alors pas nécessairement liées à une déshydratation mais peuvent être dues à une hypovolémie [5] temporaire. C'est pourquoi dans le cas d'une UFH ou UFT élevée, l'hypotension et l'hypotension orthostatique ne sont pas considérées comme des signes de déshydratation.

$PS \rightarrow Pb$, $Hb \rightarrow Pb$ **:** Une prise de poids importante peut s'expliquer soit par un poids de débranchement trop bas (*poids sec* trop bas) soit par une hyperhydratation au branchement. Inversement une prise de poids faible entre deux séances peut être due à un poids de sortie trop élevé (*poids sec* trop haut).

$PS \rightarrow Pd$, $Hd \rightarrow Pd$ **:** L'observation du poids de débranchement et l'estimation de l'hydratation vont nous donner une indication sur l'adéquation du *poids sec*. Lorsque le *poids sec* est bon, le patient aura tendance à être hyperhydraté s'il sort au dessus de son *poids sec* et déshydraté s'il sort en dessous de son *poids sec*. A l'inverse, si le patient sort en dessous de son *poids sec* sans présenter de signes de déshydratation (Hd normale), cela signifie que le *poids sec* peut être baissé. Enfin, si le patient finit la séance à son *poids sec* tout en présentant des signes de déshydratation, cela signifie que le *poids sec* est trop bas.

$PS \rightarrow Hb$, $PS \rightarrow Hd$ **:** Ces deux liens traduisent le fait que la prescription médicale a une action sur l'hydratation. Par définition, un *poids sec* trop bas va entraîner un état de déshydratation et un *poids sec* trop haut va entraîner un état d'hyperhydratation. On ne pourra donc pas conclure à un *poids sec* trop élevé si le patient est déshydraté ou a un *poids sec* trop bas si le patient est hyperhydraté.

3.3.3 Modèle dynamique

Il est souvent nécessaire de prendre en considération les faits passés pour faire un diagnostic pertinent sur l'état courant. Nous allons donc étendre le modèle de la figure 3.4 à un réseau bayésien dynamique (voir chapitre 2) pour représenter les connaissances que nous avons sur l'évolution dans le temps du processus afin de pouvoir faire apparaître des tendances dans l'évolution de l'état de santé du patient. Cette nécessité de considérer la dynamique du processus pour mettre en évidence l'état du patient est liée à l'observabilité partielle du problème qui fait que le passé nous apporte une information complémentaire pour évaluer le présent.

Nous introduisons une variable de commande Ac qui représente l'action éventuelle de changement du *poids sec* par le médecin. L'état initial du réseau dynamique B_0 est défini par le modèle statique donné par la figure 3.4. Le réseau temporel $B_{\rightarrow}$ est donné par le graphe dirigé de la figure 3.5. Pour prendre en compte la dynamique de la variable *poids sec* (PS), nous appliquons l'hypothèse markovienne d'ordre un. L'estimation de l'adéquation du *poids sec* doit se baser sur

5. Baisse du volume sanguin.

les observations de la séance, sur les éventuelles actions du médecin et sur l'estimation précédente de l'adéquation du *poids sec*. Les nœuds représentant l'état d'hydratation du patient (Hb, Hd) ne sont pas reliés d'une séance à l'autre. En effet, contrairement à la variable *poids sec*, la variable hydratation est relativement sujette à des variations journalières dont les causes ne sont pas prises en compte par notre modèle comme par exemple les repas, la température ambiante ou les efforts fournis par le patient. Nous ne faisons donc pas d'hypothèse quant à l'évolution de ces deux variables.

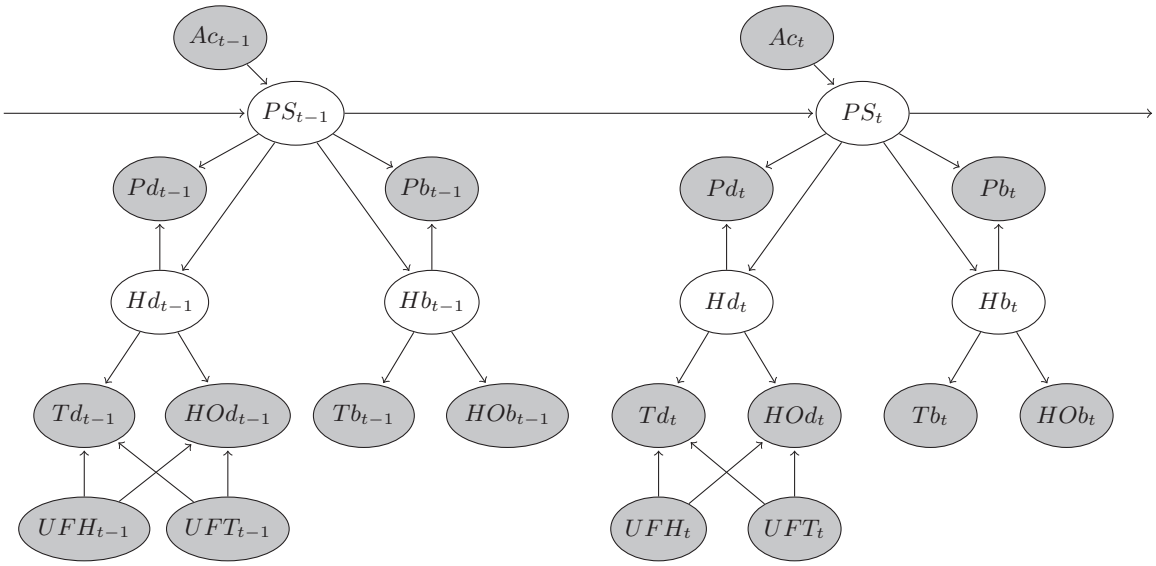


FIGURE 3.5 – Modèle de réseau bayésien dynamique pour le suivi de la qualité du *poids sec*.

3.3.4 Paramètres du réseau bayésien

Les probabilités conditionnelles associées à chaque variable du réseau ont été définies manuellement d'après l'expertise des médecins et affinées au cours de séances de travail par l'étude d'un ensemble de cas significatifs.

Ces paramètres sont adaptés au patient « moyen », c'est à dire pour lequel les règles classiques de diagnostic s'appliquent. Certaines pathologies comme l'insuffisance cardiaque modifient l'interprétation qui doit être faite des différents capteurs. Par exemple, chez un patient insuffisant cardiaque l'hyperhydratation n'entraîne pas d'hypertension car la capacité de pompe du cœur est limitée. Les patients atteints de ces pathologies ne sont, par conséquent, pas correctement suivis par ce modèle générique. Différents profils de patients peuvent être définis. D'un point de vu pratique, la notion de profil peut être intégrée au réseau bayésien par l'ajout d'une variable discrète conditionnant l'ensemble des observations. La sélection du profil se fait alors simplement en insérant une *evidence* sur la variable. Remarquons qu'avec suffisamment de données et un modèle précis des dynamiques associées à chaque profil, celui-ci pourrait être estimé par inférence.

3.4 Prétraitement des données recueillies par les capteurs

Les règles de diagnostic utilisées par le néphrologue reposent sur une interprétation des mesures numériques qui débouche sur une information symbolique. Par exemple, la mesure de la prise de poids est interprétée pour conclure ou non à la présence d'une prise de poids importante entre deux séances consécutives. De la même façon, les mesures de pressions artérielles sont intégrées pour conclure à la présence ou non d'un syndrome d'hypertension artérielle. Le raisonnement symbolique décrit par le médecin se représente bien à l'aide de variables discrètes. Le passage de la mesure numérique à l'information symbolique est plus ou moins bien défini selon les paramètres. Par exemple la notion d'hypotension orthostatique a une définition bien précise. A l'inverse la notion de prise de poids importante dépend de chaque patient et de ses variations habituelles. Nous désignons par prétraitement des données, cette étape consistant à évaluer sous la forme de probabilités le degré de confiance attribué aux différentes hypothèses symboliques à partir des valeurs numériques. Ce prétraitement est extérieur au modèle stochastique et résulte de connaissances expertes.

Nous avons utilisé, pour l'ensemble des observations du modèle, des variables discrètes comportant en général trois valeurs : *petite*, *normale* et *grande*. Ainsi, pour la mesure de tension au branchement, le prétraitement consiste à calculer les probabilités associées aux différentes valeurs de la variable Tb en fonction des mesures de pressions systoliques et diastoliques [6] :

$$\begin{aligned} P(Tb = &\ basse &&| Systole_{0...t}, Diastole_{0...t}),\\ P(Tb = &\ normale &&| Systole_{0...t}, Diastole_{0...t}),\\ P(Tb = &\ haute &&| Systole_{0...t}, Diastole_{0...t}). \end{aligned}$$

Cette phase de prétraitement peut se décrire en trois étapes répétées pour chaque capteur : le centrage des mesures, la normalisation et la discrétisation.

3.4.1 Centrage des mesures

La première étape de prétraitement consiste à centrer les mesures autour d'une valeur considérée comme normale et qui dépend de chaque patient. Cette valeur est calculée pour chaque capteur (poids, tension, ultrafiltration) d'après l'historique des mesures chez le patient et en fonction de critères médicaux donnés par l'expert.

Le poids

Le poids au branchement (Pb) : Nous nous intéressons en fait à la prise de poids journalière, c'est-à-dire la différence entre le poids au branchement de la séance (n) et le poids au débranchement de la séance précédente $(n-1)$ divisée par le nombre de jours entre les deux séances.

$$Prise\_de\_poids_{jour}(n) = \frac{Poids_{Bt}(n) - Poids_{Dbt}(n-1)}{\delta_{jour}} \tag{3.2}$$

6. La tension artérielle est donnée par 2 chiffres, la pression systolique ou maxima correspondant à la phase de contraction du cœur et la pression diastolique ou minima mesurée pendant la phase de décontraction du cœur.

La prise de poids journalière est ensuite centrée par rapport à la moyenne des prises de poids sur les 6 séances précédentes.

$$Pb(n) = Prise\_de\_poids_{jour}(n) - \frac{\sum_{k=n-6}^{n-1} Prise\_de\_poids_{jour}(k)}{6} \tag{3.3}$$

Le poids au débranchement (Pd) : Le poids de fin de séance doit normalement être très proche du *poids sec* fixé par le médecin si la prescription est suivie. Ce capteur mesure l'écart entre le poids de fin de séance et le *poids sec* prescrit.

$$Pd(n) = Poids_{Dbt}(n) - PS_{prescrit}(n) \tag{3.4}$$

La tension

Tension au branchement (Tb) : La valeur utilisée est la tension systolique corrigée en fonction de la tension pulsée et centrée autour d'une valeur de référence qui est la tension systolique supposée normale pour le patient. Cette valeur de référence est basée sur le calcul de la moyenne des tensions mesurées lors des 12 séances précédentes. Les valeurs utilisées pour le calcul de la moyenne sont saturées à 11 et à 14 de façon à ce que la valeur de référence reste dans la zone médicalement normale.

$$T_{Bt_{Pulsee}} = Syst_{Bt} - Diast_{Bt} \tag{3.5}$$

$$Corr_{Bt} = T_{Bt_{Pulsee}} - 6 \tag{3.6}$$

$$T_{Bt_{Corr}} = \begin{cases} Syst_{Bt} - Corr_{Bt} & \text{si } Syst_{Bt} > 14 \text{ et } Corr_{Bt} > 0, \\ Syst_{Bt} & \text{sinon.} \end{cases} \tag{3.7}$$

$$T_{Bt_{Sat}} = \begin{cases} 11 & \text{si } T_{Bt_{Corr}} < 11, \\ 14 & \text{si } T_{Bt_{Corr}} > 14, \\ T_{Bt_{Corr}} & \text{sinon.} \end{cases} \tag{3.8}$$

$$Tb(n) = T_{Bt_{Corr}}(n) - \frac{\sum_{k=n-12}^{n-1} T_{Bt_{Sat}}(k)}{12} \tag{3.9}$$

Tension au débranchement (Td) : Le calcul est le même que pour la tension au branchement.

$$Tb(n) = T_{Dbt_{Corr}}(n) - \frac{\sum_{k=n-12}^{n-1} T_{Dbt_{Sat}}(k)}{12} \tag{3.10}$$

Hypotension orthostatique au branchement (HOb) : Ce capteur mesure la présence ou non d'une hypotension orthostatique avant la séance. Il peut prendre les valeurs 0 ou 1.

$$HOb = \begin{cases} 1 & \text{si } SC_{Bt} - SD_{Bt} \geq 2 \text{ ou } DC_{Bt} - DD_{Bt} \geq 1, \\ 0 & \text{sinon.} \end{cases} \tag{3.11}$$

où

SC : systolique couché,
SD : systolique debout,
DC : diastolique couché,
DD : diastolique debout.

Hypotension orthostatique après la séance (HOd) : C'est l'analogue du capteur précédent pour les tensions mesurées en fin de séance.

$$HOd = \begin{cases} 1 & \text{si } SC_{Dbt} - SD_{Dbt} \geq 2 \text{ ou } DC_{Dbt} - DD_{Dbt} \geq 1, \\ 0 & \text{sinon.} \end{cases} \tag{3.12}$$

L'ultrafiltration

Ultrafiltration horaire (UFH) : La valeur de référence utilisée pour le centrage de l'ultrafiltration horaire est $PoidsSec/100$.

$$UFH = \frac{Poids_{Bt} + Apport - PoidsSec}{\delta t} - \frac{PoidsSec}{100} \tag{3.13}$$

L'apport représente la quantité de nourriture, mesurée en grammes, consommée par le patient pendant la séance. La valeur normale d'UFH est définie à un pour cent du *poids sec*, ce qui explique la fraction $PoidsSec/100$.

Ultrafiltration totale (UFT) : La valeur maximale théorique à ne pas dépasser pour l'ultrafiltration totale est $PoidsSec/10$. Nous prenons comme valeur de référence pour le centrage de ce capteur $PoidsSec/20$.

$$UFT = Poids_{Bt} + Apport - Poids_{Dbt} - \frac{PoidsSec}{20} \tag{3.14}$$

3.4.2 Normalisation

Les valeurs normalisées sont comprises entre -1 et $+1$. La normalisation fait intervenir une tolérance et une saturation. Le facteur de tolérance est introduit afin de ne pas trop prendre en compte les variations très faibles et le bruit issu des capteurs. Il permet de définir une zone normale autour de chaque valeur normale. En dehors de la zone normale, les valeurs sont saturées. Cela correspond à des vraisemblance $P(X = basse)$ ou $P(X = haute)$ égales à 1.

La déviation du capteur par rapport à la valeur normale (poids idéal, tension moyenne,...) est définie par :

$$dev = \frac{x - base}{tolerance} \tag{3.15}$$

où x est la mesure issue du capteur, $base$ est la valeur normale (prise de poids moyenne, tension moyenne,...) et $tolerance$ est le facteur de tolérance associé au capteur (voir tableau 3.2).

3.4.3 Discrétisation

La phase de discrétisation consiste à associer à une mesure continue une distribution de probabilités sur une variable aléatoire discrète. Dans notre application, les variables prennent

Capteur	Pb	Pd	Tb	Td	UFH	UFT
tolerance	1kg	0.8kg	20mmHg	20mmHg	0.6 l/h	PS/20kg

TABLE 3.2 – Les valeurs de tolérance associées à chaque capteur.

trois valeurs : *basse*, *normale*, *haute*. Le calcul des probabilités associées à chacun des états de la variable est réalisé de manière floue [Jeanpierre, 2002] par des fonctions de filtrage (voir figure 3.6).

– Une fonction sigmoïde décroissante est utilisée pour la valeur *basse*. Cette fonction est obtenue grâce au polynôme de degré 3 suivant :

$$sigmo^{-}(x) = \begin{cases} 1 & \text{si } x < -1, \\ 0 & \text{si } x > 0, \\ x^2(3+2x) & \text{sinon} \end{cases} \tag{3.16}$$

– Une fonction sigmoïde croissante est utilisée pour la valeur *haute* :

$$sigmo^{+}(x) = \begin{cases} 1 & \text{si } x < 0, \\ 0 & \text{si } x > 1, \\ x^2(3-2x) & \text{sinon} \end{cases} \tag{3.17}$$

– La valeur *normale* est obtenu en prenant le complément à 1 des deux fonctions précédentes. Cette fonction est obtenue grâce à :

$$norm(x) = 1 - (sigmo^{-}(x) + sigmo^{+}(x)) \tag{3.18}$$

La figure 3.6 donne un aperçu de ces trois opérateurs flous et montre leur complémentarité puisqu'ils couvrent l'ensemble des valeurs possibles.

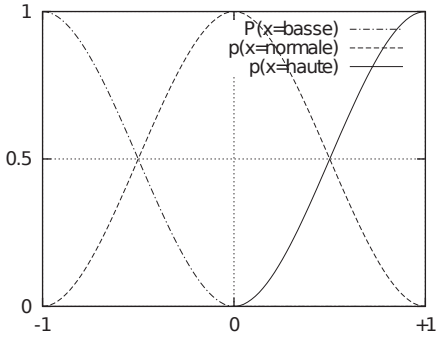


FIGURE 3.6 – Filtres flous utilisés pour le calcul des vraisemblances.

Les capteurs concernant l'hypotension orthostatique avant et après la séance, l'ultrafiltration horaire et l'ultrafiltration totale sont associés à des variables binaires ne comptant que les valeurs *normale* et *haute*.

Les valeurs de vraisemblance $P(X = basse)$, $P(X = normale)$ et $P(X = haute)$ pour les différentes variables observées sont ensuite introduites comme *evidences* incertaines (voir paragraphe 1.2.2) dans le réseau bayésien.

3.5 Interprétation des probabilités *a posteriori*

Le diagnostic que nous présentons au médecin est directement tiré de la distribution de probabilités *a posteriori* sur les valeurs de la variable *Poids Sec*. Nous utilisons l'algorithme d'inférence exacte JLO dans le DBN déroulé (voir paragraphe 2.2.3) pour propager les *evidences* des trois dernières séances et calculer

$$P(PS_t \mid seance_{t-2,\ldots,t}). \tag{3.19}$$

Ces probabilités sont présentées au médecin sous la forme d'une simple courbe

$$C_t = P(PS_t = haut \mid seance_{t-2,\ldots,t}) \tag{3.20}$$

évoluant entre -1 et $+1$. La zone positive du graphique est associée à des valeurs de *poids sec* trop grandes et la zone négative à des valeurs de *poids sec* trop petites.

L'objectif premier du système expert étant de filtrer les données afin de faciliter la tâche d'analyse du médecin, nous avons défini un mécanisme d'alerte basé sur un seuil. Lorsque les probabilités $P(PS = haut)$ ou $P(PS = bas)$ sont supérieures à 0.5 pour deux séances consécutives, la séance est ajoutée au mail d'alertes quotidien que reçoit le médecin l'avertissant ainsi d'une probable dérive du *poids sec* pour le patient.

3.6 Expérimentation

Le système est intégré à une plate-forme de la société Diatélic, partenaire de ce projet, qui recueille les données médicales et permet de présenter à la fois les données de séances et les diagnostics du système expert au médecin par l'intermédiaire d'une interface web. Les utilisateurs de l'application n'ont besoin que d'un navigateur web pour consulter les courbes de suivi et de diagnostic.

Une expérimentation de plus d'un an à été menée dans quatre centres de dialyse de la région Lorraine. Des postes clients dans les centres de dialyse ont alimenté un serveur avec les données des séances de dialyses de plus de cents patients.

La figure 3.7 montre une courbe de suivi du *poids sec* sur une période de plus de trois mois pour un patient traité par hémodialyse. Sur cette représentations sont superposées la courbe de diagnostic, les alertes générées et les actions prises par les néphrologues. Les points de la partie inférieure et supérieure du graphique correspondent respectivement à des alertes de *poids sec* trop bas et trop haut. Les flèches représentent les changements de *poids sec* décidés par le médecin. Sur cet exemple, on remarque que plusieurs alertes de *poids sec* trop haut (points de la partie supérieure) ont été générées. La validité de ces alertes a été confirmée par le médecin qui a en effet décidé de diminuer le *poids sec* à plusieurs reprises durant cette période (flèches descendantes). Une alerte de *poids sec* trop bas a été générée (point de la partie inférieure). A la séance suivante le médecin a décidé d'augmenter le *poids sec* de ce patient (flèche montante).

3.7 Conclusion sur l'application

Nous avons mis en œuvre un système à base de réseaux bayésiens dynamiques qui aide le néphrologue à suivre le *poids sec* de malades insuffisants rénaux traités par hémodialyse. Le

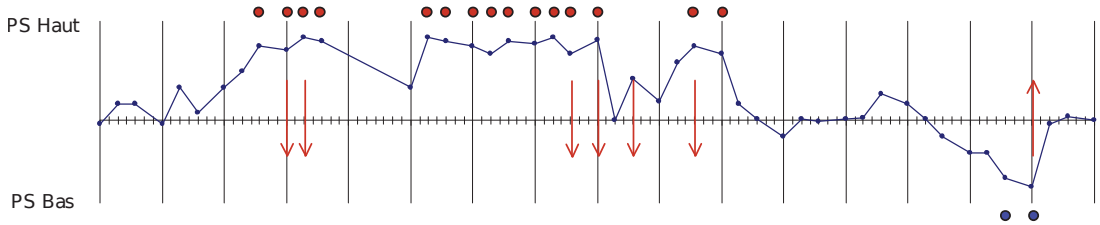


FIGURE 3.7 – Exemple de suivi du *poids sec* avec le réseau bayésien dynamique.

diagnostic donné par le système est obtenu à partir d'un ensemble restreint de données qui sont principalement le suivi du poids et des pressions artérielles du patient.

L'évaluation objective du système est très difficile étant donné que le *poids sec* réel du patient reste inaccessible. Le seul critère objectif disponible pour l'évaluation du système de télémédecine est l'évolution des indicateurs de qualité des soins comme, par exemple, le taux d'hypertension. Il est cependant très difficile de distinguer dans une telle évaluation la contribution du système d'alerte dans le système de télémédecine. Le retour que nous avons eu des médecins sur le système est positif. Le système de diagnostic est, d'après les néphrologues, fiable et concorde le plus souvent avec leurs décisions. Le système est aujourd'hui commercialisé par la société Diatélic et trouve son intérêt en particulier dans les endroits où les transports sont assez difficiles, comme à l'île de la Réunion. Cependant, l'estimation du *poids sec* reste un problème non résolu et le jeu de paramètres utilisés ici est assez réduit. Les mesures de pression artérielle sont fortement variables et dépendent d'éléments non pris en compte comme la prise d'un traitement hypotenseur, ou un changement dans l'alimentation du patient. Il serait donc intéressant d'augmenter le jeu des données prises en compte. Les machines de dialyses fournissent une information qu'il serait possible d'intégrer au système afin d'affiner les diagnostics.

3.8 Conclusions méthodologiques

L'approche utilisée nous a permis d'élaborer un système de diagnostic sans utiliser de base de données étiquetées. Le formalisme utilisé a servi de support à la discussion avec les médecins, nous permettant de représenter directement, sous forme stochastique, les règles de diagnostic permettant l'interprétation des données.

Nous pouvons distinguer deux étages dans le processus de traitement automatique des données, tous deux issus de l'intégration de la connaissance des experts. Le premier est l'étage de traitement des mesures quantitatives en vue de faire ressortir une information pertinente. Il permet notamment, pour chaque patient, de mettre en perspective les mesures journalières avec l'historique des mesures passées. Ce traitement *ad hoc* est réalisé à l'extérieur du modèle stochastique. Les résultats ainsi obtenus sont ensuite interprétés à l'aide de filtres flous pour produire une information symbolique sous forme d'*evidences* incertaines. Le second étage concerne le traitement de l'information symbolique, représentée par des variables discrètes. Cette information symbolique est manipulée dans le modèle stochastique pour évaluer la probabilité des différentes hypothèses connaissant les mesures passées. Remarquons que les filtres flous utilisés en entrée peuvent être assimilés à des lois conditionnelles. Le modèle peut donc être considéré comme un réseau hybride dans lequel seules les variables observées sont continues.

L'approche est transposable à d'autres domaines. Nous avons ainsi pu, au travers d'une autre

expérience, aborder de la même façon le problème de la prévention des rejets et de l'adaptation du traitement chez des patients transplantés. Cependant, la principale limite de cette approche est le temps d'expert qu'elle nécessite pour sa mise en œuvre et par conséquence la relative rigidité du résultat. Il faut en effet pour adapter le modèle ou lui ajouter un paramètre en entrée revisiter l'ensemble du système afin d'assurer sa cohérence. Cette expérience nous montre donc la nécessité de se tourner vers l'apprentissage automatique afin de construire des systèmes plus adaptables. La partie III de ce mémoire est consacrée à la mise en œuvre de ces techniques dans le cadre d'applications médicales.

Chapitre 4

Mesurer les paramètres de la marche

Sommaire

L'incertitude et la complexité sont deux problèmes fréquemment rencontrés lorsque l'on traite un problème du monde réel. L'incertitude provient essentiellement des observations sur le système qui sont soit partielles soit bruitées. La complexité est souvent liée au processus lui même, qui pour être complètement décrit, nécessite de nombreuses dimensions. Le problème de l'inférence consiste alors à identifier, dans un contexte incertain, une trajectoire dans un espace de très grande taille.

Nous montrerons dans ce chapitre au travers d'une application particulière qu'est la mesure des paramètres de la marche, que le formalisme des DBNs peut nous aider à raisonner sur la structure du problème et à tirer parti de cette structure pour réaliser l'inférence de manière plus efficace. Dans cet exemple les observations proviennent de caméras vidéos dont les images sont intégrées comme *evidences* dans un modèle continu. Nous proposons dans ce chapitre un algorithme d'inférence « gourmand » permettant de tirer parti de la représentation factorisée du problème afin de réduire le nombre de particules utilisées. Cet algorithme est adapté au

formalisme des DBNs et n'est par conséquent pas spécifique à l'application développée dans ce chapitre.

4.1 Problématique applicative

Dans le cadre de la télésurveillance et du maintien à domicile de personnes en perte d'autonomie, l'analyse en situation naturelle encore appelée en conditions écologiques [Paysant, 2006] est fondamentale pour estimer le degré d'autonomie de la personne, la qualité de son équilibre statique et dynamique ou encore avoir un retour sur l'impact d'une action thérapeutique. La dégradation de la marche chez les personnes âgées est un facteur important de perte d'autonomie ayant pour conséquence une aggravation du risque de chute.

L'installation de caméras vidéos utilisées comme instruments de mesure à visée médicale est envisagée en raison de leur capacité à acquérir une information complète sur les situations rencontrées par les personnes dans leur environnement quotidien. Le caractère intrusif des caméras est bien évidemment à prendre en compte et c'est pourquoi le système envisagé ne sort pas d'images mais un ensemble de mesures caractérisant la marche de la personne comme les longueurs de pas, ou les temps d'appui.

Ainsi le problème posé est ici d'estimer à chaque instant la position dans l'espace de la personne et en particulier de ses jambes à partir des images provenant d'une ou plusieurs caméras.

4.2 Capture du mouvement sans marqueurs

Il existe des systèmes commerciaux très précis pour réaliser la capture de mouvements 3D. Ces systèmes supposent l'utilisation de marqueurs posés sur différentes articulations du corps, ce qui limite leur utilisation à des situations de laboratoire et ne nous permet pas d'envisager d'utiliser ces systèmes au domicile des personnes.

Travailler sans marqueurs rend bien évidemment l'observation de la scène plus difficile. La plupart des travaux visant à reconstruire globalement le mouvement d'un sujet dans une scène consiste à formuler des hypothèses, soit à partir d'un modèle humanoïde 3D ([Deutscher *et al.*, 2000], [Saboune, 2008]), soit à partir d'un ensemble d'exemples pré-définis, et à confronter ces hypothèses aux observations provenant des caméras vidéos. L'évaluation des hypothèses peut se faire en utilisant une mesure de ressemblance entre les silhouettes ou les contours provenant d'une part de l'image réelle et, d'autre part, de l'hypothèse.

Ce problème se formalise bien comme un problème d'inférence bayésienne en considérant comme vecteur d'état X_t la configuration de la personne à l'instant t et comme observations Y_t les images provenant des caméras. La mesure de ressemblance permet pour une configuration particulière $\mathbf{x}_t$ et une observation particulière $\mathbf{y}_t$ d'évaluer la probabilité conditionnelle $P(Y_t = \mathbf{y}_t \mid X_t = \mathbf{x}_t)$. La mesure de ressemblance entre la forme de l'hypothèse et l'image observée est utilisée pour mettre à jour la probabilité que nous attribuons à cette hypothèse comme illustré par la figure 4.1. La probabilité mesure bien ici un degré de confiance dans la vérité d'une proposition. Le domaine des hypothèses est continu. La configuration de l'humanoïde est décrite dans l'espace par un vecteur dont la dimension est définie par le nombre de degrés de libertés du modèle 3D (une modélisation réaliste du corps humain nécessite 31 degrés de liberté). La fonction

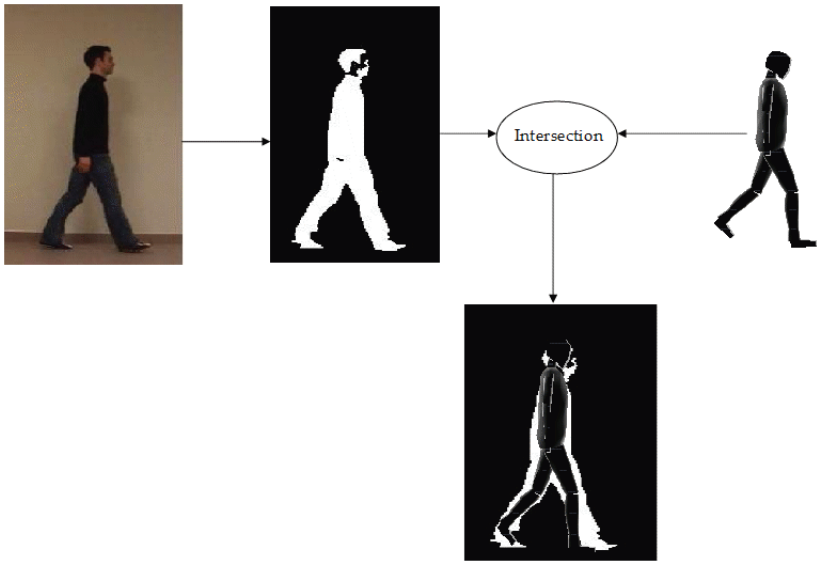


FIGURE 4.1 – Evaluation de la ressemblance entre l'image observée et une hypothèse. Cette ressemblance est interprétée comme la probabilité $P(Y_t = \mathbf{y}_t \mid X_t = \mathbf{x}_t)$ d'observer l'image sachant l'hypothèse.

d'observation est également continue. Elle prend en entrée un vecteur de pixels et il est bien sûr très difficile de présager de sa forme. Nous pouvons cependant affirmer assez facilement que celle-ci est multimodale. En considérant par exemple une vue de profil, il est assez facile de voir que deux hypothèses différentes auront une probabilité élevée, la première avec la jambe droite devant, la seconde avec jambe gauche devant. Plus généralement les symétries du corps humain donnent lieu à des ambiguïtés qui se traduisent par une observation fortement multimodale.

Le filtrage particulaire (voir paragraphe 2.2.5) permet de réaliser l'inférence approchée quelque soit la forme des distributions de probabilités. Celui-ci s'applique donc tout naturellement au problème de l'estimation des mouvements de marche.

4.3 Réduire le nombre de particules

Dans le problème posé, nous ne cherchons pas à estimer l'intégralité de la distribution *a posteriori*. Seul le maximum *a posteriori* (MAP) nous intéresse. L'enjeu est donc de pouvoir suivre dans le temps la position du MAP avec un minimum de particules. Pour un nombre de particules donné, le choix des zones de la loi *a posteriori* à échantillonner soulève le dilemme classique entre exploration et exploitation. Utiliser plus de particules pour l'exploration permet de s'assurer que le maximum trouvé est bien un maximum global, alors qu'échantillonner la loi *a posteriori* autour du maximum trouvé permet d'affiner la solution. Une heuristique classique consiste à ré-échantillonner régulièrement la distribution en prenant un nombre de particules proportionnel à la densité de probabilité estimée. Cependant le nombre de particules requis pour estimer efficacement le MAP grandit exponentiellement avec le nombre de dimensions du problème [McCormick and Isard, 2000].

Différentes approches ont été proposées pour réduire le nombre de particules nécessaires au suivi 3D des mouvements d'une personne. Dans [Deutscher *et al.*, 2000], les auteurs ont uti-

lisé une technique de recuit simulé pour trouver le maximum global et affiner itérativement la solution. Saboune et Charpillet [Saboune and Charpillet, 2005] ont proposé une exploration partiellement déterministe de l'espace d'états pour maintenir les efforts d'exploration avec peu de particules. L'approche consiste à échantillonner régulièrement sur un intervalle la loi de transition pour certaines dimensions du problème qui dans le cadre applicatif sont considérées comme plus importantes. Cet échantillonnage régulier a pour effet de favoriser une exploration efficace, les particules étant réparties uniformément sur les intervalles considérés. Dans ce travail, certaines configurations supposées fréquentes sont testées à chaque itération par l'utilisation de particules dites « statiques » injectées dans la population à chaque pas de temps.

4.4 Approche factorisée

Si l'on demande à un humain de rechercher la position d'une personne en jouant sur la position et la configuration d'un pantin articulé de manière à faire coïncider les deux images, il est très probable que l'humain commencera par positionner le torse avant de s'intéresser successivement aux différents segments articulés formant les jambes et les bras. Le formalisme des réseaux bayésiens nous permet de représenter les dépendances en chaîne, à l'origine de cette intuition, que nous avons, qu'il est préférable de rechercher la position du torse avant de s'attaquer à celle de la cuisse et qu'il est plus efficace de positionner la cuisse avant de rechercher la position de la jambe. L'approche que nous proposons est inspirée par une démarche intellectuelle humaine tout en étant fondée sur les propriétés mathématiques du problème.

4.4.1 Factorisation de l'espace d'états

Le corps humain peut être considéré comme une chaîne cinématique ouverte, c'est à dire un ensemble d'articulations connectant un ensemble de segments $\{S^i\}_{i=0}^N$ de telle sorte que tout segment soit connecté à chacun des autres par un et un seul chemin. L'ensemble de la chaîne a donc la forme d'un arbre dont la racine peut être choisie arbitrairement parmi l'un des segments. Soit $\mathcal{X} = \{X^0, \ldots, X^1\}$ la configuration de la chaîne, décrite par les configurations 3D X^i de chaque segment S^i numéroté par ordre topologique. En notant p_i l'indice du segment parent de S^i (c'est à dire le premier segment du chemin liant le segment S^i à la racine), nous pouvons formuler l'hypothèse d'indépendance suivante : *la configuration X^i du segment S^i est indépendante de la configuration des segments précédents de la chaîne (côté racine) connaissant la position X^{p_i} du segment parent.* Cette indépendance conditionnelle peut s'illustrer dans le cas du corps humain en remarquant que la position de l'avant-bras est indépendante de la position du torse connaissant la position du bras. Il en découle une représentation factorisée de la distribution de probabilité sur la configuration de la chaîne [3] :

$$P(\mathcal{X}) = \prod_{i=1}^{N} P(X^i \mid X^{p_i}). \tag{4.1}$$

Ainsi la loi de probabilité sur la configuration d'une chaîne cinématique peut être décrite par un réseau bayésien dont la structure découle directement de celle de la chaîne.

La loi conditionnelle $P(X^i \mid X^{p_i})$ décrit la relation existant entre deux segments consécutifs. Elle dépend des contraintes de l'articulation connectant les deux segments. En remarquant que la position absolue X^i est une fonction de X^{p_i} et de la configuration R^i de l'articulation

reliant X^i à X^{p_i} ($X^i = f_i(X^{p_i}, R^i)$), nous pouvons utiliser une représentation plus compacte de la configuration de la chaîne en considérant la position absolue du segment racine X^0 les positions relatives $R^1, \ldots, R^N$ des autres segments. La factorisation de l'ensemble de variables $\mathcal{R} = (X^0, R^1, \ldots, R^N)$ s'écrit alors :

$$P(X^0, R^1, \ldots, R^N) = P(X^0) \prod_{i=1}^{N} P(R^i). \tag{4.2}$$

Cette description ne tient compte que des contraintes physiques sur les éléments de la chaîne, et ne considère pas le processus à l'origine du positionnement de ces éléments.

Nous supposerons que le contrôle des différentes articulations résulte de processus indépendants. Cette hypothèse est fausse en pratique : les mouvements de marche reposent sur une coordination des mouvements des différents membres (mouvement alterné des jambes et balancement synchronisés des bras). Cependant cette approximation est justifiée si l'on ne cherche pas à modéliser le processus à l'origine des mouvements, comme c'est le cas dans notre cadre applicatif, puisque nous souhaitons justement pouvoir détecter des marches anormales et donc ne pas présupposer que les mouvements sont normaux. Ainsi la factorisation proposée reste valide dans une approche dynamique sous l'hypothèse d'indépendance des contrôles.

Pour décrire le processus $\mathcal{R}_t = \{X_t^0, R_t^1 \ldots, R_t^N\}$, en nous plaçant dans un cadre markovien, il faut bien sûr s'intéresser à la configuration $\mathcal{R}_{t-1}$ de la chaîne à l'instant précédent. Les processus de contrôle des différentes articulations étant supposés indépendants nous pouvons décrire la dynamique du système comme le produit des lois dynamiques de chaque articulation :

$$P(\mathcal{R}_t \mid \mathcal{R}_{t-1}) = P(X_t^0 \mid X_{t-1}^0) \prod_{i=1}^{N} P(R_t^i \mid R_{t-1}^i). \tag{4.3}$$

Puisque $X_t^i = f_i(X_t^{p_i}, R_t^i)$, $(X_t^{p_i}, R_t^i)$ d-sépare X_t^i des variables précédentes. Et la loi jointe de $(\mathcal{X}_t, \mathcal{R}_t)$ s'écrit :

$$P(\mathcal{X}_t, \mathcal{R}_t \mid \mathcal{X}_{t-1}, \mathcal{R}_{t-1}) = P(X_t^0 \mid X_{t-1}^0) \prod_{i=1}^{N} P(R_t^i \mid R_{t-1}^i) \prod_{i=1}^{N} P(X_t^i \mid X^{p_i}, R_t^i) \tag{4.4}$$

avec

$$P(X_t^i \mid X_t^{p_i}, R_t^i) = \delta(X_t^i, f_i(X_t^{p_i}, R_t^i)) \tag{4.5}$$

où $\delta(x, f_i(X_t^{p_i}, R_t^i))$ est la distribution de Dirac centrée en $f_i(X_t^{p_i}, R_t^i)$. La figure 4.2 illustre cette factorisation.

4.4.2 Fonction d'observation

L'approche proposée par Saboune [Saboune, 2008], inspirée notamment des travaux de Deutscher [Deutscher *et al.*, 2000], consiste à évaluer, par filtrage particulaire, la vraisemblance d'une configuration en comparant la silhouette d'un modèle 3D de la personne suivie à celle provenant des cameras. La silhouette de la personne suivie est extraite sur chaque caméra par soustraction du fond et seuillage. Le modèle 3D est ensuite projeté sur chaque plan caméra dans la configuration à évaluer. Le poids des particules est calculé en fonction de la surface de la zone de

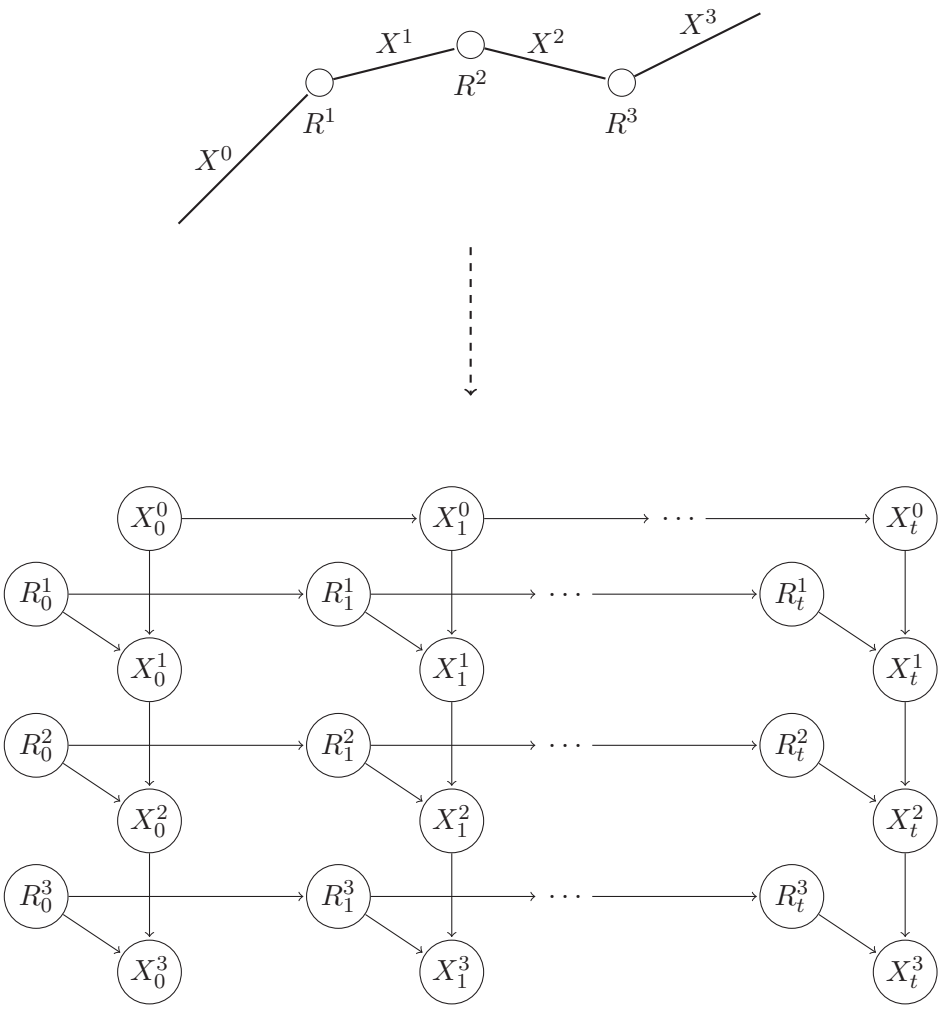


FIGURE 4.2 – Représentation graphique de la factorisation d'une chaîne cinématique.

FIGURE 4.3 – Illustration de la décomposition de la fonction d'observation en une séquence d'évaluations locales. La zone bleue représente les pixels communs masqués lors de l'évaluation des précédents segments. Les points de cette zone ne sont pas considérés dans le comptage des points d'intersection N_c, ni dans le calcul de N_s et N_m.

recouvrement entre la silhouette extraite et le modèle projeté (voir figure 4.1). Le poids w d'une configuration évalué sur une caméra est donné par :

$$w = \frac{N_c}{N_s + N_m} \tag{4.6}$$

où

- N_c est le nombre de pixels communs entre la silhouette extraite et l'image synthétique,
- N_s est le nombre de pixels de la silhouette extraite n'appartenant pas à l'image synthétique,
- N_m est le nombre de pixels de l'image synthétique n'appartenant pas à la silhouette extraite.

Dans cette approche, la vraisemblance d'une configuration est évaluée de manière globale. Ainsi la fonction d'observation dépend de l'intégralité du vecteur d'état. En utilisant la variable Y pour désigner l'observation, le poids d'une configuration s'écrit :

$$w_t = P(y \mid X_t^0, \ldots, X_t^N). \tag{4.7}$$

Comme nous le verrons dans l'algorithme 2, il est nécessaire de pouvoir évaluer une hypothèse partielle pour tirer parti de la factorisation du vecteur d'état. Nous proposons donc de décomposer la fonction d'observation en un ensemble de mesures locales à chaque segment. Cette décomposition peut être faite en projetant les segments individuellement sur la silhouette extraite. Afin de ne pas compter plusieurs fois une même zone de la silhouette dans les points communs pour différents segments, nous proposons de masquer les intersections obtenues lors de l'évaluation d'un segment avant d'évaluer le segment suivant. Les pixels masqués de la silhouette n'interviendront pas dans la détermination des valeurs de N_c, N_s et N_m (voir figure 4.3).

La fonction d'observation que nous proposons est composée d'une chaîne d'observations locales, chacune dépendant de la précédente à cause du processus de masquage des pixels communs. Un ordre dans l'évaluation des segments doit donc être défini. En notant Y^i l'observation associée au segment i, et $pre(Y^i)$ l'observation précédente, la fonction globale s'écrit :

$$P(Y_t \mid \mathcal{X}_t) = \prod_{i=1}^{N} P(Y_t^i \mid X_t^i, pre(Y_t^i)) \tag{4.8}$$

La représentation graphique correspondante est donnée par la figure 4.4.

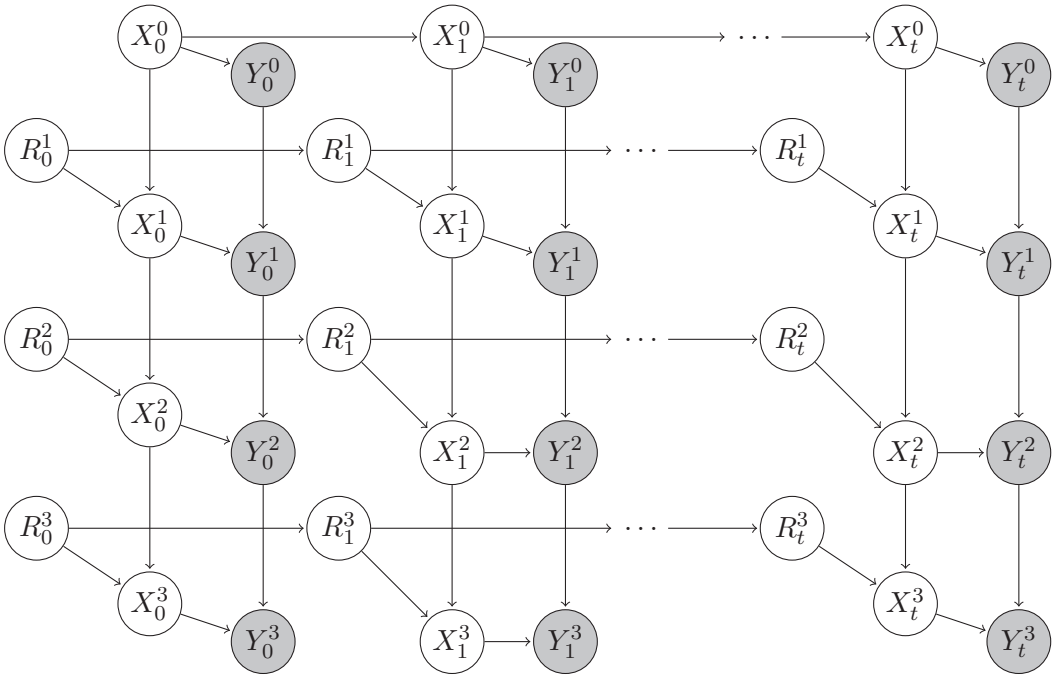


FIGURE 4.4 – Représentation graphique de la factorisation d'une chaîne cinématique et de la fonction d'observation.

4.5 Algorithme gourmand

L'algorithme 1, présenté dans le paragraphe 2.2.5, permet de générer les particules à partir des paramètres du DBN et d'intégrer les observations en mettant à jour les poids w^i associés à chaque configuration. Dans l'algorithme de référence, le test de rééchantillonnage de la distribution peut avoir lieu au plus une fois à chaque pas de temps en fonction du nombre de particules efficaces. Etant donné que nous ne nous intéressons pas à l'ensemble de la distribution mais au suivi du maximum *a posteriori*, nous proposons d'utiliser une variante « gourmande » [3] décrite par l'algorithme 2. Dans cette variante, le test de rééchantillonnage est fait après l'introduction de chaque observation. Cette modification permet d'éliminer tôt les hypothèses les moins probables afin de recentrer les efforts d'exploration autour des hypothèses les plus vraisemblables. A partir d'un ensemble $\{z_{t-1}^i, w_{t-1}^i\}_{i=1}^N$ de particules pondérées à l'instant $t-1$, l'algorithme consiste à générer un ensemble de particules pondérées $\{z_t^i, w_t^i\}_{i=1}^N$ à l'instant t. En partant d'un ensemble de vecteurs vides $\{z_t^i\}_{i=1}^N$ et en parcourant le graphe du DBN par ordre topologique :

- si le nœud rencontré correspond à une variable cachée, une valeur de la variable Z_t^k est tirée pour chaque particule z_t^i en suivant la distribution $P(Z_t^k \mid Pa(Z_{t-1}^k = u))$ où u est la valeur de $Pa(Z_{t-1}^k)$ dans la particule z_t^i ou sa particule parente z_{t-1}^i (nous rappelons que $Pa(Z_t^k) \in (\mathcal{Z}_t, \mathcal{Z}_{t-1})$) ;
- sinon le nœud rencontré est une observation qui permet de mettre à jour le poids des particules avec

$$w_t^i = w_t^i \times P(Z_t^k = z_t^i \mid Pa(Z_t^k) = u) \tag{4.9}$$

et la distribution peut alors être éventuellement rééchantillonnée comme illustré par la figure 4.5.

Entrées : Une distribution échantillonnée $\{z_{t-1}^i, w_{t-1}^i\}_{i=1}^N$ à l'instant $t-1$ et un vecteur d'observations y_t à l'instant t
pour $i = 1$ *à* N **faire**
  $w_t^i = w_{t-1}^i$;
  z_t^i = vecteur vide de longueur N;
fin
pour *Chaque nœud* k *par ordre topologique* **faire**
  si Z_t^k *n'est pas observée* **alors**
    pour $i = 1$ *à* N **faire**
      Soit u la valeur de $Pa(Z_t^k)$ dans (z_{t-1}^i, z_t^i) ;
      Compléter le vecteur z_t^i avec une valeur tirée selon la loi $P(Z_t^k \mid Pa(Z_t^k) = u)$.;
    fin
  sinon
    pour $i = 1$ *à* N **faire**
      Soit u la valeur de $Pa(Z_t^k)$ dans (z_{t-1}^i, z_t^i) ;
      Compléter le vecteur z_t^i avec la valeur de Z_t^k dans y_t;
      $w_t^i = w_t^i \times P(Z_t^k = z_t^i \mid Pa(Z_t^k) = u$;
    fin
    Calculer $w_t = \sum_i w_t^i$;
    Normaliser les w_t^i : $w_t^i = \frac{w_t^i}{w_t}$;
    Calculer $N_{eff} = \frac{1}{\sum_i (w_t^i)^2}$;
    si $N_{eff} < seuil$ **alors**
      Sélectionner les particules selon leur poids;
      Uniformiser les poids : $w_t^i = \frac{1}{N}$;
    fin
  fin
fin

Algorithme 2: Algorithme d'inférence gourmand, avec rééchantillonnage éventuel de la distribution après chaque observation. Dans cet algorithme Z_t^k représente la $k^{\text{ième}}$ variable du DBN : $\mathcal{Z}_t = (Z_t^1, \ldots, Z^K)$. z_t^i et w_t^i désignent respectivement la configuration $(\mathcal{Z}_t = z_t^i)$ et le poids de la particule i à l'instant t.

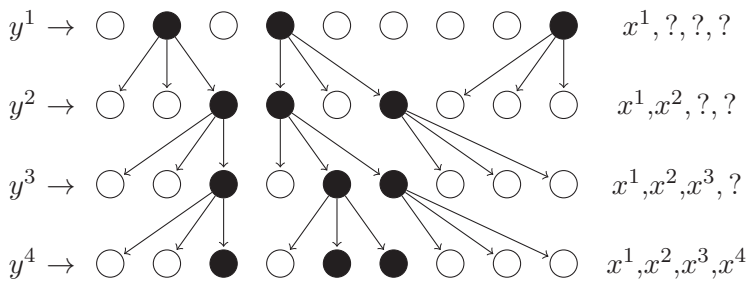


FIGURE 4.5 – Illustration du rééchantillonnage de la distribution après l'introduction de chaque observation. Ici les 3 configurations les plus probables sont conservées et dupliquées 3 fois avant d'être complétées.

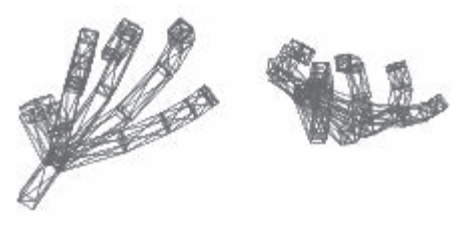


FIGURE 4.6 – Modèle 3D formant une chaîne cinématique à 15 degrés de liberté.

Comme nous le verrons dans le paragraphe 4.6. L'approche proposée peut être combinée avec la méthode d'échantillonnage déterministe de Saboune et al. [Saboune and Charpillet, 2005]. Rien empêche en effet d'échantillonner régulièrement la loi $P(Z_t^k \mid Pa(Z_t^k) = u)$ au moment de générer les particules. De la même façon, le critère de décision pour le rééchantillonnage de la distribution peut être adapté pour, par exemple, décider d'effectuer le rééchantillonnage après chaque variable observée.

4.6 Mouvements de main simulés

Nous présentons ici une évaluation expérimentale de l'algorithme sur des mouvements de main simulés. Le choix d'évaluer la méthode en simulation dans un premier temps, nous permet de nous affranchir des contraintes liées à l'expérimentation réelle telles que les problèmes de calibration des caméras et de pré-traitement des images (extraction des ombres, extraction de la silhouette). Ce type d'évaluation présente également l'avantage de nous permettre de disposer d'une référence précise concernant les trajectoires suivies. En revanche, il n'est pas possible de transposer les résultats à un cadre applicatif réel, les qualités de robustesse au bruit n'étant pas évaluées en simulation.

Le modèle utilisé contient 15 articulations pour un total de 15 degrés de liberté (voir figure 4.6). La seule hypothèse faite sur la dynamique du système est l'existence de contraintes sur la vitesse angulaire de chaque articulation. Les distributions conditionnelles correspondantes sont donc des lois uniformes sur un intervalle centré autour de la position angulaire de l'articulation à l'instant précédent. Nous avons combiné l'algorithme 2 avec l'échantillonnage déterministe de [Saboune and Charpillet, 2005] en tirant, après chaque observation, 5 fois les 5 particules les plus probables, ce qui revient à dire que chacune des 5 particules sélectionnées est dupliquée 5 fois. La dimension suivante est ensuite complétée en prenant 5 valeurs réparties uniformément dans l'intervalle couvert par la loi conditionnelle. Etant donné que cette façon de rééchantillonner la distribution ne modifie pas le poids relatif des différentes hypothèses les poids des particules ne sont pas uniformisés. Conserver le poids des particules après le rééchantillonnage nous permet de maintenir la proportion $\alpha_t(x_t) \propto \sum_{i=1}^{N} w_t^i \delta(x_t, x_t^i)$ (voir paragraphe 2.2.5). Nous travaillons donc avec une population de 25 particules.

La scène est enregistrée depuis différents points de vus par 3 caméras virtuelles. Le contrôle du modèle 3D est réalisé en utilisant plusieurs lois sinusoïdales déphasées. 25 particules suffisent pour obtenir un suivi visuellement satisfaisant. Remarquons que l'appréciation visuelle, même si celle-ci paraît subjective, permet une critique efficace des mouvements reconstruits, l'œil percevant assez bien les irrégularités ou saccades dans les mouvements. Nous avons mesuré l'erreur relative des mouvements reconstruits en divisant la distance euclidienne entre la vraie position 3D des

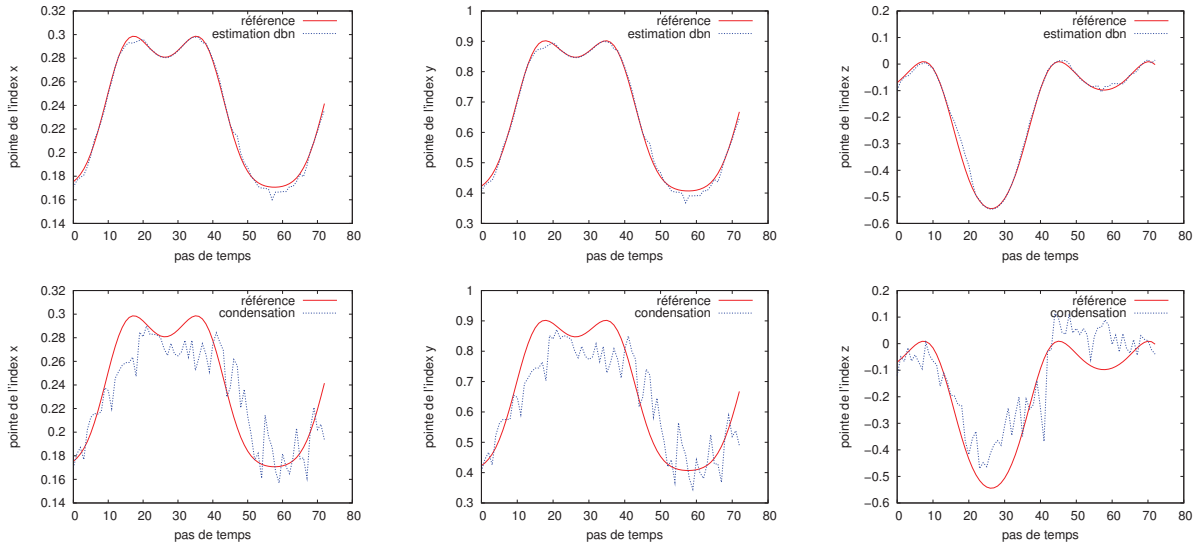


FIGURE 4.7 – Comparaison entre la trajectoire réelle et estimée de la pointe de l'index d'une part en utilisant l'approche factorisée (en haut) et d'autre part avec l'algorithme condensation (en bas).

articulations et la position 3D estimée par l'amplitude des mouvements. L'erreur moyenne ainsi calculée sur l'ensemble des articulations et sur l'ensemble de la séquence est de 4%. A titre de comparaison, l'algorithme condensation fonctionnant avec 375 particules produit une erreur de 13%. Le choix d'utiliser 375 particules pour l'algorithme condensation provient du fait que dans notre approche les 25 particules sont réévaluées pour chaque degré de liberté. Ainsi, même si, dans notre approche, les estimations sont des estimations locales, nous effectuons $15 \times 25 = 375$ comparaisons. Il est donc équitable d'utiliser 375 particules dans condensation pour comparer les algorithmes. La figure 4.7 représente d'un coté des trajectoires reconstruites avec l'approche factorisée et, de l'autre, les résultats obtenu avec l'algorithme *condensation*. Même si ce dernier réussit à suivre le maximum global malgré le petit nombre de particules utilisées, il apparaît clairement que l'approche factorisée permet d'obtenir une reconstruction beaucoup plus lisse à nombre égal de particules. Au niveau de la pointe de l'index l'erreur relative moyenne est de 2% pour l'algorithme factorisé contre 19% pour condensation.

4.7 Application au suivi des paramètres de marche

Comme mentionné précédemment, le corps humain peut être représenté par une chaîne cinématique ouverte comportant 5 branches attachées au torse : une pour chaque membre et une cinquième pour la tête. Cette représentation, bien que simpliste (la colonne vertébrale est ici considérée comme rigide) permet de suivre les mouvements de marche. La figure 4.8 illustre un tel modèle.

La factorisation du vecteur d'état permet de réduire le nombre de particules nécessaires au suivi de la personne. En utilisant la même méthode de diffusion et de sélection des particules que dans l'algorithme IPF [Saboune and Charpillet, 2005] et le modèle dynamique, nous obtenons des résultats visuels similaires avec 2000 évaluations de particules que IPF utilisant 6000

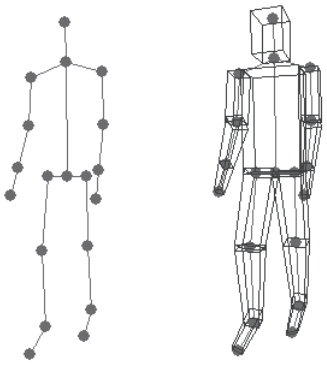


FIGURE 4.8 – Un modèle 3D articulé du corps humain, résumé à 19 articulations pour un total de 31 degrés de liberté.

particules. Rappelons qu'il s'agit d'un modèle à 31 degrés de liberté et que seules les contraintes en position et en vitesse angulaire sont considérées. Ainsi, pour chaque articulation, la position angulaire à l'instant t est choisie dans un intervalle autour de la position angulaire à l'instant $t-1$ dont la largeur dépend des limites physiologiques de l'articulation. Dans l'expérimentation comparative, un sujet marchant sur une trajectoires rectiligne était filmé par deux caméras vidéos synchronisées, fonctionnant à 25 images par seconde et fournissant une résolution de 720×580 points par image.

L'utilisation du formalisme des DBN permet également de faciliter la fusion de plusieurs sources d'information. Il suffit pour cela d'ajouter des variables observées conditionnées par les variables d'état.

4.7.1 Observations multiples

La précision de la reconstruction dépend du nombre de particules utilisées mais aussi de la quantité d'information utilisée pour faire l'estimation. Il est bien sûr possible de réduire les ambiguïtés liées aux observations en augmentant le nombre de caméras. Ceci permet de diminuer l'importance des phénomènes d'occlusion, comme par exemple le masquage d'une jambe par l'autre, ou le masquage de la personne suivie par un objet de l'environnement. Augmenter la quantité d'information induit nécessairement un coût de calcul supplémentaire lors de l'évaluation des différentes particules, mais ceci nous permet également de diminuer le nombre de particules nécessaires au suivi.

Silhouette

Nous reprenons la fonction d'observation présentée au paragraphe 4.4.2. Les segments sont projetés de façon séquentielle sur la silhouette et le décompte des points de silhouette non encore masqués N_c est divisé par la somme des points $(N_s + N_m)$ en dehors de la zone commune :

$$P = \frac{N_c}{N_s + N_m} \tag{4.10}$$

Ainsi, en nommant $P_t^{l,k}$ l'observation correspondant au segment l faite sur la camera k à l'instant t, nous calculons la vraisemblance de la particule i à l'aide d'une distribution softmax :

$$P(P_t^{l,k}|z_t^i) = \frac{\exp\left(\frac{P_i}{\tau_{P_l}}\right)}{\sum_j \exp\left(\frac{P_j}{\tau_{P_l}}\right)}. \tag{4.11}$$

où P_i est le décompte réalisé pour la particule i et τ_{P_l} est une *température* définie empiriquement pour le segment l. Remarquons que le dénominateur n'a pas vraiment d'importance, si ce n'est pour jouer le rôle de facteur d'échelle, puisqu'il ne modifie pas l'importance relative donnée aux différentes hypothèses.

Image de contours

Les contours peuvent être extraits de l'image réelle par l'utilisation d'un filtre, comme le filtre de Sobel [Sobel and Feldman, 1968], permettant de calculer le gradient de l'intensité en chaque point. Si la norme du gradient dépasse un certain seuil, on considérera qu'il s'agit d'un point de contour. L'image de contour apporte une information complémentaire à celle de la silhouette. Celle-ci est couramment utilisée en reconnaissance de formes [Gavrila, 1996]. Une distance peut être calculée entre le contour extrait et un contour de référence [Barrow *et al.*, 1977]. Cette mesure nous permet de définir une nouvelle observation C_t à fusionner avec l'image de silhouette. Le calcul de la distance entre les contours repose sur la création d'une carte de distances aux dimensions de l'image réelle et sur laquelle chaque point contient la distance en pixels du point de contour le plus proche. Cette carte de distance, appelée en anglais *chamfer map*, se construit avec l'algorithme 3. Nous utilisons pour mesurer la ressemblance entre le contour d'un segment du modèle et l'observation le critère suivant :

$$C = -\frac{1}{n}\sum_{k=1}^{n} d_k \tag{4.12}$$

où les d_k représentent les n valeurs de la carte de distance correspondant aux n points du contour du segment. La figure 4.9 illustre le calcul de distance à partir de l'image de contour.

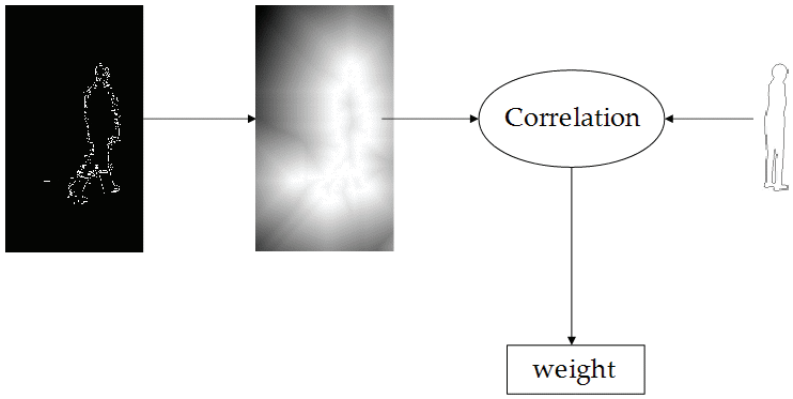


FIGURE 4.9 – Calcul de la ressemblance entre deux images de contour.

Nous intégrons la vraisemblance $P(C_t^{l,k}|z_t^i)$ de la particule z_t^i, calculée pour le segment l sur

Input : Image binaire de contour I $(N \times M)$
Output : Carte de distance D $(N \times M)$
`/* Initialisation */`
pour $i = 1$ *to* N **faire**
  pour $j = 1$ *to* M **faire**
    si $I(i,j)$ *est un point de contour* **alors**
      $D(i,j) \leftarrow 0$
    fin
    $D(i,j) \leftarrow D_{max}$
  fin
fin
`/* Passe avant */`
pour $i = 1$ *to* N **faire**
  pour $j = 1$ *to* M **faire**
    $D[i][j] \leftarrow min(D[i][j], D[i][j-1] + 2, D[i-1][j] + 2,$
    $D[i-1][j-1] + 3, D[i-1][j+1] + 3)$
  fin
fin
`/* Passe arrière */`
pour $i = N$ *to* 1 **faire**
  pour $j = M$ *to* 1 **faire**
    $D[i][j] \leftarrow min(D[i][j], D[i+1][j] + 2, D[i][j+1] + 2,$
    $D[i+1][j-1] + 3, D[i+1][j+1] + 3)$
  fin
fin

Algorithme 3: Construction de la carte de distance

la caméra k, à l'aide d'une distribution softmax :

$$P(C_t^{l,k}|z_t^i) = \frac{\exp\left(\frac{C_i}{\tau_C}\right)}{\sum_j \exp\left(\frac{C_j}{\tau_C}\right)}. \tag{4.13}$$

où τ_C est une *température* définie empiriquement.

Direction de marche

L'évaluation de l'orientation de la personne suivie est assez difficile étant donné que le corps humain est essentiellement vertical et présente donc une silhouette relativement homogène par rotation autour de l'axe vertical. La dynamique du modèle ne tient pas compte de l'existence d'un lien entre l'orientation de la personne et le sens de la marche. S'agissant de mouvements de marche naturels, il est en effet probable que la personne soit orientée dans le sens de sa marche, ou éventuellement, dans certains cas particuliers, dans le sens opposé (marche à reculons, personne s'asseyant sur une chaise, ...). Concernant la dynamique de l'orientation, nous avons deux hypothèses à fusionner. D'une part, la continuité du mouvement imposant que l'orientation à l'instant $(t+1)$ se trouve dans un certain voisinage de l'orientation à l'instant (t) et, d'autre part, l'hypothèse d'une ressemblance *probable* entre l'orientation de la trajectoire suivie et celle de la personne. Nous proposons d'intégrer cette seconde hypothèse sous la forme d'une observation O_t, calculée en estimant le cap de la trajectoire suivie et conditionnée par l'orientation de la personne (voir figure 4.10). Le terme observation est ici abusif puisque la valeur de la variable que nous introduisons ne sera pas estimée à partir d'informations extérieures, mais à partir de l'estimation de la position de la personne dans le passé.

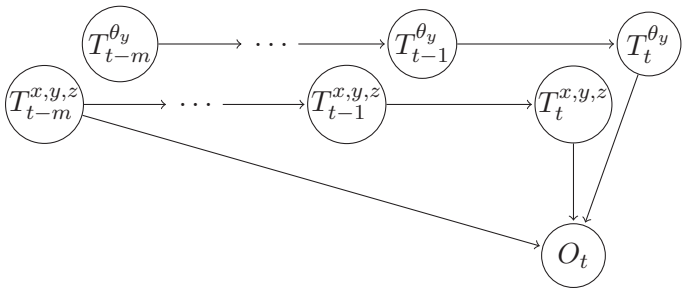


FIGURE 4.10 – Pseudo-observation de l'orientation de la personne. L'estimation faite dépend de la position passée, de l'observation courante.

En notant $T_t^{x,y,z}$ la position du torse estimée à l'instant t et en remarquant que plus la vitesse de déplacement est importante plus nous pouvons avoir confiance dans notre estimation de l'orientation du vecteur de marche, nous proposons d'utiliser la mesure de vraisemblance suivante (voir la figure 4.11) :

$$P(O_t \mid z_t^i, z_{t-m}^i) = \frac{\exp\left(\frac{f(z_t^i, z_{t-m}^i)}{\tau_O}\right)}{\sum_j \exp\left(\frac{f(z_t^j, z_{t-m}^j)}{\tau_O}\right)} \tag{4.14}$$

où

$$f(z_t^i, z_{t-m}^i) = -(O_t - T_t^{\theta_y})^2 d(T_t^{x,y,z}, T_{t-m}^{x,y,z}). \tag{4.15}$$

Dans l'équation 4.15 $d(T_t^{x,y,z}, T_{t-m}^{x,y,z})$ est la distance euclidienne entre les positions à $(t-m)$ et à (t) :

$$d(T_t^{x,y,z}, T_{t-m}^{x,y,z}) = \sqrt{(T_t^{x,y,z} - T_{t-m}^{x,y,z})'(T_t^{x,y,z} - T_{t-m}^{x,y,z})} \tag{4.16}$$

O_t désigne l'orientation de la trajectoire du torse (voir figure 4.11) et τ_O est une température.

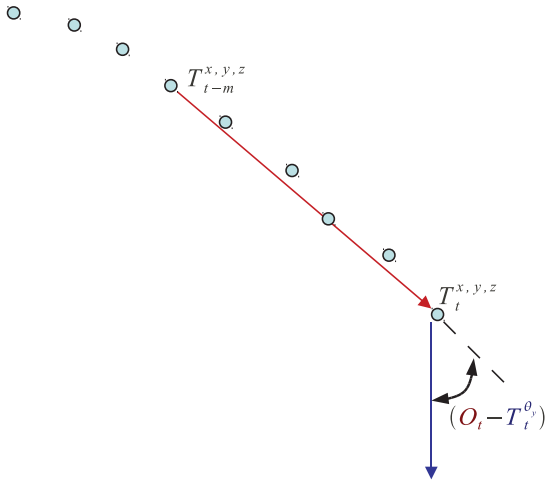


FIGURE 4.11 – Pseudo-observation de l'orientation de la personne en fonction de la trajectoire de marche.

Fusion des vraisemblances

Nous avons donc, pour une variable particulière du modèle, un certain nombre d'observations dépendant en particulier du nombre de caméras. Pour fusionner ces observations nous faisons les hypothèses suivantes :

1. nous considérons que les différentes mesures obtenues sur une image particulière sont indépendantes connaissant le vecteur d'état, ce qui nous permet d'écrire

$$P(P_{l,t}^k, C_{l,t}^k \mid z_t^i) = P(P_{l,t}^k \mid z_t^i) P(C_{l,t}^k \mid z_t^i); \tag{4.17}$$

2. nous faisons, de la même façon, l'hypothèse que les mesures obtenues sur les différentes caméras sont indépendantes connaissant le vecteur d'état

$$P(P_{l,t}^1, \ldots, P_{l,t}^1 \mid z_t^i) = \prod_k P(P_{l,t}^k \mid z_t^i). \tag{4.18}$$

Ainsi la mise à jour du poids de la particule i après visite de la variable cachée correspondant à la position du segment l est donnée par

$$w_i = w_i \times \prod_k P(P_{l,t}^k \mid z_t^i) \prod_k P(C_{l,t}^k \mid z_t^i). \tag{4.19}$$

La pseudo-observation de l'orientation du torse est intégrée après visite de la variable T^{θ_y} par

$$w_i = w_i \times P(O_t \mid z_t^i, z_{t-m}^i). \tag{4.20}$$

La figure 4.12 synthétise les relations entre les différentes observations et les différentes composantes du vecteur d'état dans le cas d'une seule caméra.

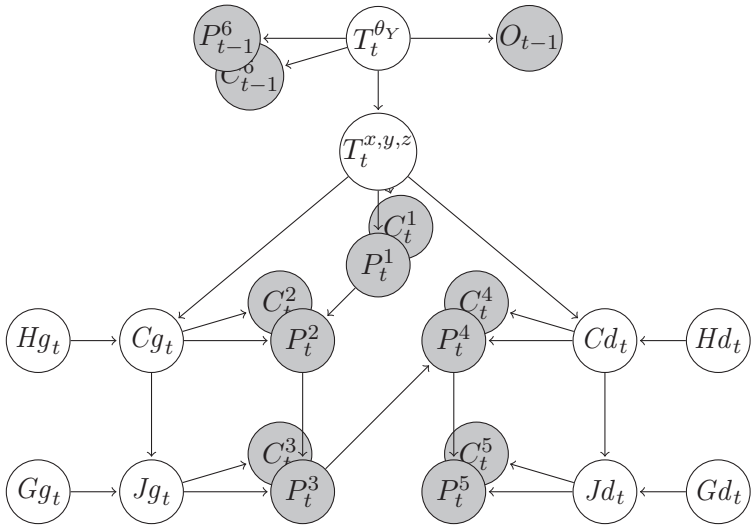


FIGURE 4.12 – Représentation d'un DBN modélisant partiellement le corps humain assimilé à une chaîne cinématique. Seules les jambes apparaissent ici. La position 3D des jambes gauche et droite (*Jg*, *Jd*) dépendent de la position angulaire des genoux (*Gg*, *Gd*) et de la position 3D des cuisses (*Cg*, *Cd*). De façon similaire, la position des cuisses peut être déterminée à partir de la position du torse T et des hanches (*Hg*, *Hd*).

Méthode de ré-échantillonnage

Nous proposons de comparer expérimentalement deux stratégies de ré-échantillonnage.

Diffusion Statique (Sd) La première consiste, comme dans l'algorithme de Saboune et al [Saboune and Charpillet, 2005], à sélectionner un nombre prédéterminé K de particules et à dupliquer chaque particule un certain nombre de fois de façon à échantillonner régulièrement les intervalles de valeurs à explorer. Dans notre cas cet échantillonnage *deterministe* s'applique après chaque introduction d'observation pendant le parcours du graphe. Le poids des particules n'est pris en compte que pour la sélection des K particules les plus probables. Afin de conserver l'information provenant du passé, les poids des particules ne doivent donc pas être uniformisés.

Diffusion Adaptative (Ad) La méthode de ré-échantillonnage classique, utilisée en particulier dans l'algorithme *Condensation*, consiste à faire un tirage des particules proportionnellement à leur poids et à échantillonner les intervalles de valeurs avec une méthode de Monte-Carlo. Dans notre cas, le modèle dynamique utilisé pour les différentes articulations fait que le tirage doit se faire avec une loi uniforme sur un intervalle autour de la position précédente. Remarquons que cette méthode de diffusion peut être combinée avec le taux de survie déterministe qui consiste à chaque ré-échantillonnage à éliminer les particules les moins probable de façon à effectuer le tirage à partir des K particules les plus probables, toujours proportionnellement à leur poids. Ce paramètre permet de favoriser plus ou moins l'exploitation face à l'exploration.

Fonctions	Moyenne	Max
PCO_Ad	8.2	15.5
PC_Ad	12.8	23.9
PO_Ad	13.2	26.5

Fonctions	Moyenne	Max
PCO_Sd	12.9	26.0
PC_Sd	13.2	25.3
PO_Sd	13.3	24.32

TABLE 4.1 – Erreur 2D moyenne et maximale commise (en pixels) sur l'estimation de la position du torse.

Fonctions	Moyenne	Max
PCO_Ad	10.1	17.1
PC_Ad	11.7	28.9
PO_Ad	11.1	19.5

Fonctions	Moyenne	Max
PCO_Sd	10.0	21.2
PC_Sd	11.3	24.6
PO_Sd	10.8	23.8

TABLE 4.2 – Erreur 2D moyenne et maximale commise (en pixels) sur l'estimation de la position du genou gauche.

Evaluation expérimentale

L'étude présentée ici [2] a été réalisée sur une séquence d'un sujet marchant sur une trajectoire circulaire et filmé par 3 caméras analogiques synchronisées fonctionnant à 25 images par secondes. Pour évaluer la contribution des différentes fonctions d'observation, nous reconstruisons une référence à partir d'un étiquetage manuel des vidéos. Pour ce faire, nous marquons image par image la position de plusieurs articulations ce qui nous permet de reconstruire la trajectoire 3D suivie par ces différents points d'intérêt. La trajectoire ainsi reconstruite est cependant loin d'être parfaite. C'est pourquoi nous nous focalisons sur l'étiquetage 2D, construit sur une échelle en pixels. La référence 3D nous permet simplement d'avoir une idée approximative sur l'ordre de grandeur de l'erreur commise par l'estimation du système.

Nous comparons les résultats obtenus d'un côté en utilisant la méthode de ré-échantillonnage proportionnelle au poids des particules (Ad) et, de l'autre, en utilisant un échantillonnage déterministe des intervalles (Sd). Dans chaque cas, nous comparons l'erreur 2D moyenne et maximale calculée sur l'ensemble de la séquence en utilisant différentes combinaisons des fonctions d'observation. Les résultats chiffrés sont présentés sur les tableaux 4.1, 4.2, 4.3 pour les erreurs commises respectivement sur la position du torse, du genou gauche et de la cheville gauche. Le tableau 4.4 donne à titre indicatif l'erreur de reconstruction dans l'espace pour la position du torse, du genou gauche et de la cheville gauche. Ces valeurs sont à rapporter aux dimensions de la scène puisque le sujet évolue à 4 mètres environ des 3 caméras.

Fonctions	Moyenne	Max
PCO_Ad	11.9	41.1
PC_Ad	16.8	63.2
PO_Ad	12.6	45.3

Fonctions	Moyenne	Max
PCO_Sd	14.7	59.4
PC_Sd	16.8	61.5
PO_Sd	14.4	62.2

TABLE 4.3 – Erreur 2D moyenne et maximale commise (en pixels) sur l'estimation de la position de la cheville gauche gauche.

Partie	Moyenne	Max
Torse	8.0	14.8
Genou gauche	9.0	19.5
Cheville gauche	11.1	39.9

TABLE 4.4 – Erreur 3D (en cm) commise lors de l'estimation en utilisant la méthode adaptative et les 3 fonctions d'observation (P,C et O).

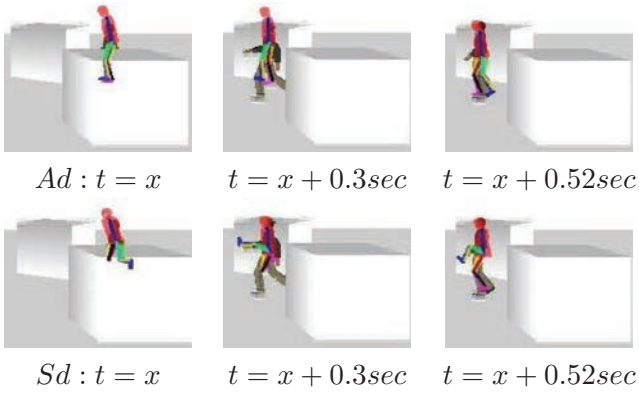


FIGURE 4.13 – Comparaison du temps de récupération entre la diffusion statique (Sd) et la diffusion adaptative (Ad).

Les résultats numériques nous confirment l'intérêt de fusionner différentes sources d'information complémentaires pour améliorer la précision du suivi quelque soit la méthode de ré-échantillonnage utilisée. Par ailleurs, l'échantillonnage déterministe donne dans notre algorithme factorisé de moins bons résultats que la méthode adaptative alors que le nombre de particules utilisées pour la méthode adaptative (50) était inférieur au nombre de particules utilisées dans la méthode déterministe (200). Il semble que la sélection d'un petit nombre prédéterminé de particules après chaque observation face perdre en robustesse vis à vis de l'incertitude liée aux observations. Il est en effet peu probable que de multiples hypothèses faites sur la position d'un segment proche de la racine, puissent survivre longtemps si le taux de sélection des particules est faible, ce qui est le cas dans la diffusion déterministe (5% dans notre expérimentation). Dans le cas de la diffusion adaptative, le taux de sélection est variable puisqu'il dépend de la valeur des poids. Cette hypothèse d'une plus grande robustesse de la diffusion adaptative face à l'incertitude de l'observation se confirme en simulation en ajoutant des objets dans la scène. Comme illustré sur la figure 4.13, le nombre d'images nécessaires à la récupération après une occultation partielle de la personne suivie est plus petit dans le cas de la diffusion adaptative (15 images soit 0,52 secondes) que dans le cas de la diffusion statique (52 images soit 2,08 secondes).

Les figures 4.14 et 4.15 montrent des exemples de suivi dans deux situations particulières. Dans la première, la personne suivie s'agenouille et dans la seconde la personne suivie effectue un mouvement rapide (saut).

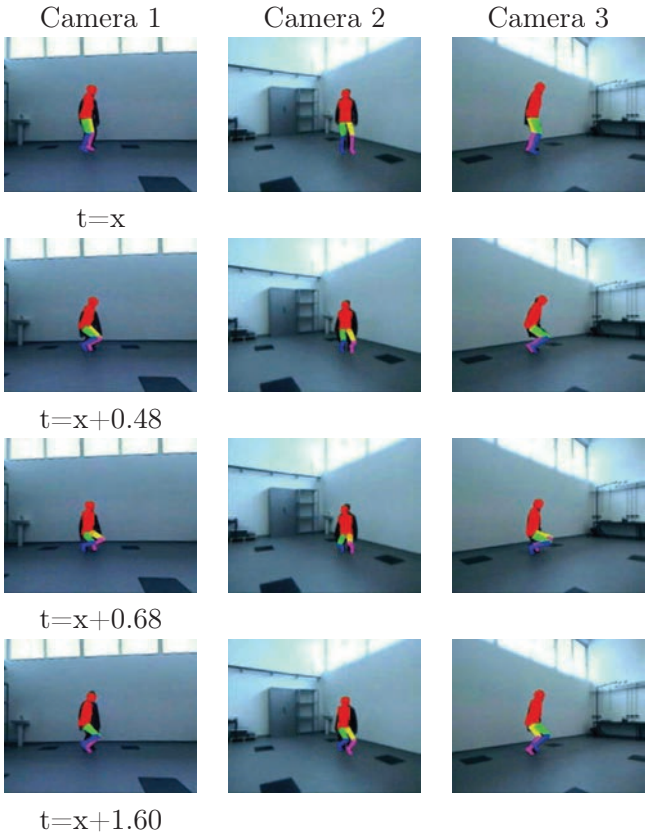


FIGURE 4.14 – Exemple de suivi d'un sujet posant les genoux au sol.

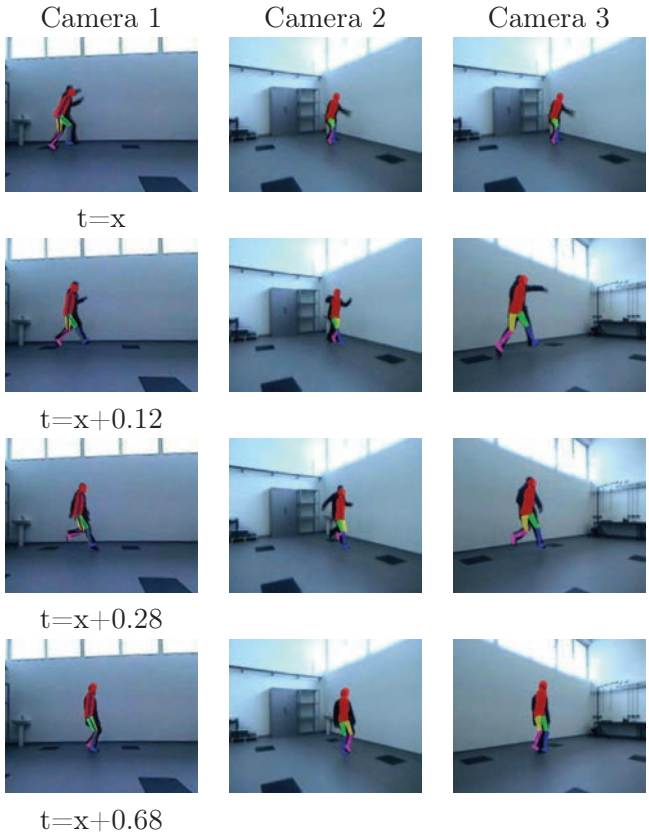


FIGURE 4.15 – Exemple de suivi d'un sujet effectuant un saut.

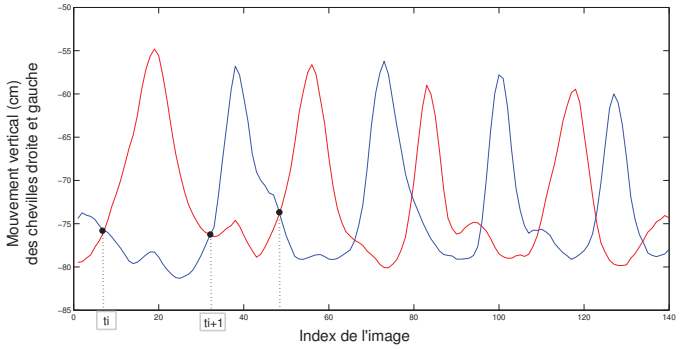


FIGURE 4.16 – Mouvement vertical estimé des chevilles droite (rouge) et gauche (bleu).

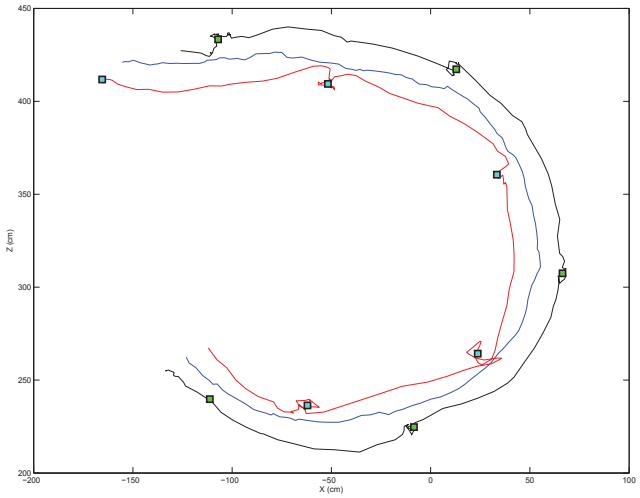


FIGURE 4.17 – Vue de la position du torse (en bleu) de la cheville gauche(en noir) et de la cheville droite (en rouge) en projection sur le sol. Les carrés représentent les positions estimées des appuis.

4.7.2 Reconstruction de la position des pas

La marche se caractérise par une succession de pas pendant lesquels au moins l'un des deux pieds est en contact avec le sol alors que le second se déplace. Le mouvement vertical de la cheville permet de mettre en évidence ce phénomène comme illustré sur la figure 4.16. La position verticale de la cheville peut être utilisée de façon simple pour reconstruire la position et la durée des pas. En effet, le pied le plus bas peut être considéré comme fixe et en contact avec le sol. Nous pouvons ainsi calculer sa position moyenne sur toute la durée de contact avec le sol afin de reconstruire la position des pas comme illustré sur la figure 4.17. Cette méthode de reconstruction des pas n'est envisageable que pour des déplacements dans un plan horizontal, ce qui exclus par exemple des montées ou descentes d'escaliers.

4.8 Evaluation en situation réaliste

Une étude statistique a été réalisée, dans le cadre d'un stage de master [Dubois, 2010], avec l'objectif d'évaluer le système développé dans des situations réalistes. Cette étude a consisté à faire marcher 7 sujets dans 13 situations différentes adaptées à la problématique du suivi de personnes âgées, comprenant des marches normales, à grands pas, à petits pas, avec ramassage d'un objet au sol, avec une canne, avec un déambulateur, avec une jupe ou une robe de chambre, avec un obstacle sur le parcours,... Le protocole suivi a consisté à marquer sur des rouleaux de papier les traces de pas grâce à des tampons imbibés d'encre et posés sous les chaussures.

L'étude a porté sur l'évaluation de la capacité du système à estimer la longueur et la durée des pas. Ces deux paramètres sont particulièrement importants car il permettent de déduire un certain nombre de variables considérées comme caractéristiques de la marche [Auvinet *et al.*, 2003]. L'irrégularité des longueurs de pas semble, en effet, être une variable pertinente pour la prédiction des chutes. Les valeurs de référence pour la longueur des pas ont été mesurées manuellement à partir des traces sur les rouleaux de papier. Un logiciel d'étiquetage de flux vidéos a été utilisé pour mesurer la durée des pas. Un premier résultat de cette étude est donc la constitution d'une base de référence permettant une évaluation objective du système et de ses évolutions futures.

Le travail a permis de montrer que le système développé répondait bien au problème de la localisation de la personne quelques soient les situations étudiées. Cette étude a également révélé certaines limites de l'approches, en particulier lorsque le système est confronté à des situations de marche assistée par un déambulateur. Les mesures estimées dans les situations de marche normale, avec une robe ou encore avec une canne ne présentent pas de différences statistiquement significatives avec les mesures réelles. Cependant, l'absence de différences avérées n'est pas une preuve d'égalité et la variabilité de l'erreur de reconstruction reste relativement importante comme l'illustre la figure 4.18. Celle-ci révèle en particulier une forte variabilité de l'erreur de reconstruction pour la situation en robe de chambre. La reconstruction de la position des pas dans l'espace n'est donc pas toujours précise. Cependant, l'étude a montré que les valeurs moyennes étaient pertinentes et nous permettaient par exemple de différencier les situations étudiées. Il est donc envisageable d'utiliser l'outil de reconstruction, dans son état actuel, pour analyser l'évolution dans le temps de mesures moyennes calculées, par exemple, sur un ou plusieurs jours.

L'estimation de la longueur des pas dépend à la fois de l'algorithme de reconstruction des mouvements et de la méthode d'extraction de la position des pas. L'extraction des pas est réalisée de manière simpliste en utilisant un paramètre unique qui est la trajectoire verticale des chevilles au cours du temps. Cette méthode est très sensible aux erreurs commises lors de la reconstruction des mouvements et nous envisageons maintenant d'utiliser un modèle bayésien pour estimer la position des pas à partir des mouvements reconstruits. L'extraction de la position des pas serait alors fondée sur la hauteur des chevilles ainsi que sur d'autres paramètres comme le mouvement du torse et le mouvement des cuisses par rapport au torse. Il existe probablement une marge de progression importante concernant la fonction d'observation. La méthode utilisée dépend de la soustraction du fond pour l'extraction de la silhouette et de la détection des contours. L'extraction du fond manque de robustesse lorsque la couleur des vêtements est proche de celle du fond. La détection des contours peut également être perturbée dans certaines situations, par exemple, par l'existence de motifs sur les vêtements. Nous envisageons donc de faire évoluer la fonction d'observation de façon à la rendre plus robuste. La texture et la couleur des vêtements constituent une information que nous pouvons apprendre ou adapter en début de séquence. Au niveau de

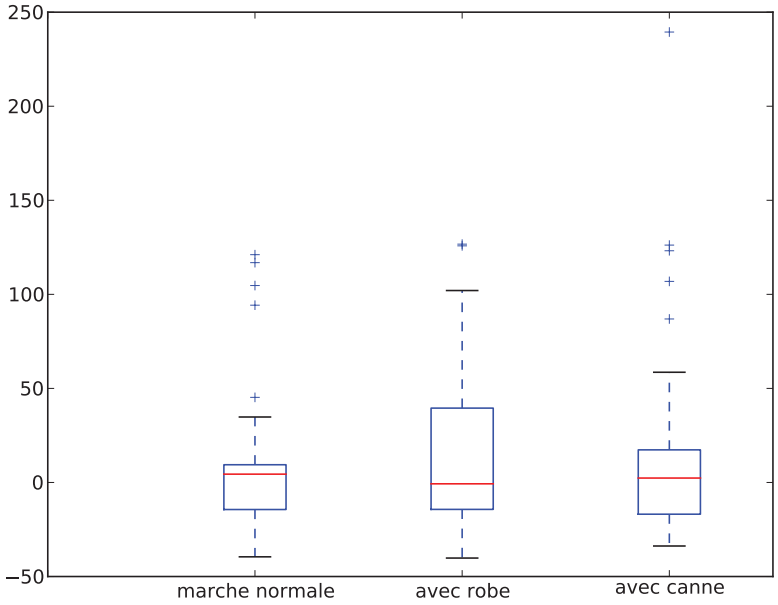


FIGURE 4.18 – Dispersion des erreurs de reconstruction de la longueur des pas, exprimées en centimètres, selon différentes situations étudiées.

l'algorithme de filtrage, certains paramètres pourront être adaptés. Les données de l'étude nous ont, par exemple, permis d'adapter le taux de survie des particules, caractérisé par le paramètre K (voir paragraphe 4.7.1), qui joue un rôle dans la balance entre les efforts d'exploration et d'exploitation. Les données de l'étude vont nous permettre d'évaluer plus objectivement l'impact du nombre de particules utilisées sur la précision de la reconstruction afin de mieux régler ce paramètre dans un compromis temps-précision.

4.9 Conclusion

La contribution est d'abord ici d'ordre méthodologique. En partant d'une formulation bayésienne du problème de la reconstruction de mouvements de marche à partir de caméras vidéos, nous avons pu formaliser une démarche intellectuelle qui consiste à tirer parti des indépendances conditionnelles issues de la structure du problème. Nous montrons ainsi que le formalisme des DBNs nous sert à la fois de support à l'expression d'un raisonnement et d'outil pour la résolution d'un problème d'inférence difficile en raison des dimensions de l'espace d'état. La contribution est également algorithmique puisque nous avons proposé un algorithme de filtrage particulaire « gourmand » permettant de tirer parti de la factorisation du problème pour sa résolution dans le cas de la recherche du maximum *a posteriori*.

Sur le plan logiciel, le travail réalisé avec Abdallah Dib [2] a permis de réduire de façon importante le temps nécessaire à l'analyse des flux vidéos. Celle-ci n'est pas encore réalisée en temps réel mais cet objectif est désormais envisageable. Ce travail s'inscrit dans un projet à moyen terme de réalisation d'un capteur ambiant de capture du mouvement destiné au suivi des

mouvements de marche. Le projet est toujours en cours et les perspectives de développement sont d'au moins deux natures. La première vise la mise en place d'un service d'alerte destiné au grand public pour aider au maintien à domicile d'un proche en perte d'autonomie. Le second développement envisagé est l'élaboration d'un capteur à visée médicale capable de mesurer les paramètres de la marche en milieu écologique ainsi que d'autres éléments d'appréciation comme le niveau d'activité.

Conclusion

Nous nous sommes intéressés dans cette seconde partie à la question de la modélisation d'une démarche médicale sous la forme d'un problème d'inférence bayésienne visant à estimer l'état d'une variable cachée représentant l'état du patient. Dans cette approche, la construction du modèle se fait à la suite d'un dialogue avec un expert humain du domaine. Dans notre expérience sur l'hémodialyse, présentée dans le chapitre 3, le formalisme graphique des réseaux bayésiens nous a servi de support à la discussion avec les néphrologues. Ce type de modélisation est particulièrement adapté au cas où il n'est pas possible ou difficile d'obtenir des données étiquetées. Dans notre exemple, le *poids sec* n'est pas une donnée mesurable et il est parfois même difficile d'obtenir un consensus des médecins quant à l'adéquation de la prescription du fait de l'interprétation variable de certaines observations. La validation objective du modèle est difficile en raison justement de l'absence de données étiquettées. Celle-ci se fait donc par des études de cas. Notre approche, qui se fonde sur la démarche plutôt que sur les conclusions, nous permet néanmoins d'établir un modèle de diagnostic dans ce contexte incertain. La principale limite est que l'élaboration du modèle nécessite un investissement important de la part de l'expert humain et qu'il est difficile de faire évoluer ou d'adapter le modèle sans que celui-ci intervienne à nouveau.

Le réseau bayésien est un support à la représentation de connaissances, qui nous permet de dialoguer avec l'expert et de modéliser sa démarche, mais ce formalisme est également un outil permettant de mettre en place un raisonnement et de réaliser l'inférence efficacement. Nous avons étudié dans le chapitre 4 un problème d'inférence particulièrement difficile, consistant à identifier une trajectoire suivie dans un espace de grande dimension. La structure même du modèle décrit une connaissance sur le problème qui peut être exploitée pour réaliser l'inférence. Nous avons ainsi pu donner une représentation factorisée de la position du corps humain permettant d'estimer plus efficacement des mouvements de marche. Sur le plan algorithmique nous avons proposé un algorithme de filtrage particulaire « groumand » permettant de tirer parti de la factorisation de problème pour l'estimation du maximum *a posteriori*.

Troisième partie

De l'exemple au modèle, l'apprenti apprenant

Introduction

Dans cette troisième partie, nous abordons le problème de l'apprentissage des paramètres du modèle dans une démarche où la transmission de la connaissance se fait par l'intermédiaire d'un ensemble d'exemples. Il s'agit de reproduire une expertise humaine à partir de données numériques.

Le chapitre 5 introduit le principe de l'identification du modèle par maximisation de la vraisemblance qui est une mesure de l'adéquation du modèle avec les observations. Nous présentons en particulier l'algorithme EM, généralisé au cas d'un DBN, qui permet de rechercher, pour une structure donnée, le modèle présentant la plus grande vraisemblance. Deux approches de l'apprentissage à partir d'exemples sont ensuite envisagées.

Dans la première approche, qui fait l'objet des chapitres 6 et 7, le problème de la télésurveillance médicale est traité comme un problème d'apprentissage supervisé qui consiste à apprendre une fonction pour laquelle un ensemble de points représentatifs des entrées et des sorties correspondantes est connu. Une limite de l'apprentissage supervisé est qu'il nécessite une grande quantité de données couvrant les différentes situations possibles. Par ailleurs, dans ce type d'approches, la connaissance modélisée est celle d'un ou plusieurs experts particuliers.

Nous envisageons finalement, dans le chapitre 8, de traiter le problème de la télésurveillance comme un problème de contrôle en cherchant à amener ou à maintenir le patient dans un certain état. Cette dernière approche est particulièrement intéressante car le but à atteindre est alors caractérisé par une fonction de récompense qui évalue l'efficacité du traitement. Celle-ci peut être issue de mesures objectives, comme par exemple des relevés biologiques. L'apprentissage permet de découvrir parmi les actions prise par l'expert, celles qui ont effectivement donné le résultat attendu. Il est ainsi possible d'apprendre une politique de meilleure qualité que celle utilisée lors de la constitution des exemples. Un autre avantage est qu'il n'est pas nécessaire de définir un modèle *a priori*.

Chapitre 5

Identification du modèle

Sommaire

L'apprentissage est dit supervisé lorsqu'il se base sur un ensemble d'exemples contenant à la fois les entrées et les sorties d'une fonction dont nous cherchons à apprendre les paramètres. Ce mode d'apprentissage est particulièrement adapté aux problèmes de classification dans lesquels il est possible pour un expert de donner les sorties correspondant à un ensemble représentatif de cas. Ce chapitre présente brièvement l'algorithme EM dans le cas d'un réseau bayésien dynamique. Il permet d'apprendre les paramètres du modèle en maximisant sa vraisemblance.

5.1 Maximisation de la vraisemblance

La vraisemblance d'un modèle $\mathcal{M}_\Theta$ pour un ensemble de données $\mathcal{D}$ est la mesure $P(\mathcal{D}|\mathcal{M}_\Theta)$. La mesure de vraisemblance permet de quantifier si le modèle explique bien les données. Cependant cette mesure ne peut être considérée en absolu et interprétée comme la probabilité que le modèle soit le bon car elle ne tient pas compte de la probabilité *a priori* du modèle. Nous avions illustré dans la partie 1.2 l'importance de la prise en compte de la loi *a priori* dans l'interprétation d'un résultat.

Dans l'apprentissage par maximisation de la vraisemblance nous faisons l'hypothèse que les

modèles évalués sont équiprobables *a priori* et donc :

$$P(\mathcal{M}_\theta \mid \mathcal{D}) = P(\mathcal{D} \mid \mathcal{M}_\theta)\frac{P(\mathcal{M}_\theta)}{P(\mathcal{D})} = \alpha P(\mathcal{D} \mid \mathcal{M}_\theta) \tag{5.1}$$

Cette hypothèse peut se justifier lorsqu'il s'agit de rechercher les valeurs des paramètres d'un modèle car nous n'avons en général aucune information permettant de quantifier la probabilité *a priori* des différents modèles. En revanche, dans le cas où l'on chercherait à comparer des modèles de tailles différentes avec plus où moins de paramètres, les vraisemblances ne sont plus comparables. Des modèles plus compliqués permettent en général d'atteindre des vraisemblances plus élevées. Cependant le rasoir d'Ockham (principe de parcimonie), souvent utilisé en apprentissage, nous invite à privilégier un modèle plus simple ce qui se traduit dans le raisonnent bayésien par utiliser un *a priori* non uniforme sur les modèles.

5.2 Observabilité totale

Etant donné un graphe dirigé composé de N variables discrètes $(X_1, \ldots, X_N)$, nous cherchons à estimer les paramètres $\theta = \{P(X_i \mid Pa(X_i))\}_{i=1,\ldots,N}$ du réseau à partir d'un ensemble d'observations $\mathcal{D}$ sur l'intégralité des N variables du réseau. Notons :

- q_i le nombre d'états de la variable i,
- r_i le nombre de combinaisons des parents $Pa(X_i)$ de la variable i,
- θ_{ijk} la valeur de la probabilité conditionnelle $P(X_i = j|Pa(X_i) = k)$.

La vraisemblance d'une observation $d_l \in \mathcal{D}$ est donnée par :

$$P(d_l|\theta) = \prod_{i=1}^{N}\prod_{i=1}^{q_i}\prod_{k=1}^{r_i}\theta_{ijk}^{\delta_{ijkl}} \tag{5.2}$$

ou δ_{ijkl} est la fonction qui vaut 1 si $X_i = j$ et $Pa(X_i) = k$ dans l'échantillon d_l et 0 sinon. Sur l'ensemble de la base

$$P(\mathcal{D}|\theta) = \prod_{l}\prod_{i=1}^{N}\prod_{i=1}^{q_i}\prod_{k=1}^{r_i}\theta_{ijk}^{\delta_{ijkl}} \tag{5.3}$$

$$P(\mathcal{D}|\theta) = \prod_{i=1}^{N}\prod_{i=1}^{q_i}\prod_{k=1}^{r_i}\theta_{ijk}^{N_{ijk}} \tag{5.4}$$

$$\log P(\mathcal{D}|\theta) = \sum_{i=1}^{N}\sum_{i=1}^{q_i}\sum_{k=1}^{r_i}N_{ijk}\log\theta_{ijk} \tag{5.5}$$

L'estimateur du maximum de vraisemblance est donné par :

$$\theta_{ijk} = \frac{N_{ijk}}{\sum_{j=1}^{q_i} N_{ijk}} \tag{5.6}$$

5.3 Observabilité partielle

Dans le cas où les vecteurs de données $\mathbf{d_l}$ seraient partiellement remplis la maximisation de la vraisemblance peut se faire avec un algorithme EM. L'étape *Expectation* consiste à estimer les

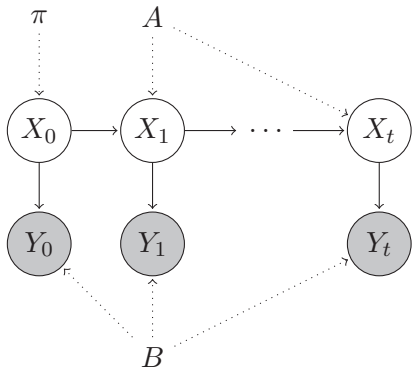


FIGURE 5.1 – HMM caractérisé par les paramètres π, A et B.

données manquantes par inférence dans le modèle courant et l'étape *Maximization* est similaire au cas complètement observable. Le décompte des nombres de cas est simplement remplacé par l'espérance du nombre de cas.

Ainsi l'estimateur du maximum de vraisemblance est donné par :

$$\theta_{ijk} = \frac{E\left[N_{ijk}\right]}{\sum_{j=1}^{q_i} E\left[N_{ijk}\right]} \tag{5.7}$$

où

$$E\left[N_{ijk}\right] = \sum_{l} P(X_i = j, Pa(X_i) = k \mid d_l) \tag{5.8}$$

est l'espérance du nombre de configurations où $X_i = j$ et $Pa(X_i) = k$ sur $\mathcal{D}$, l'ensemble des observations.

En pratique, l'étape *Expectation* consiste pour chaque vecteur $\mathbf{d_l}$ à poser les *evidences* correspondant à chaque variable observée puis à réaliser l'inférence pour estimer les données manquantes. Rappelons que, dans l'algorithme de l'arbre de jonctions (voir paragraphe 1.3.2), l'étape de moralisation nous assure qu'il existe pour chaque variable au moins une clique contenant à la fois la variable et ses parents. Après inférence, nous pouvons donc extraire la probabilité $P(X_i = j, Pa(X_i) = k \mid d_l)$ à partir du potentiel de cette clique. Les paramètres θ_{ijk} sont ensuite ré-estimés d'après l'équation 5.7 pour maximiser la vraisemblance du modèle.

5.4 Apprentissage des paramètres d'un DBN

Comme présentés au chapitre 2, les modèles dynamiques auxquels nous nous intéressons sont des modèles du premier ordre caractérisant un processus markovien. L'hypothèse markovienne nous permet, pour l'apprentissage de la dynamique du processus, de considérer l'ensemble des transitions observées indépendamment les unes des autres. Autrement dit, pour une séquence observée $(x_0, x_1, \ldots, x_T)$, l'apprentissage de la dynamique $P(X_t|X_{t-1})$ pourra se faire à partir de l'ensemble des transitions $\{(x_{i-1}, x_i\}_{i=1,\ldots,T}$.

5.4.1 Cas particulier d'un HMM : algorithme Baum-Welch

L'algorithme Baum-Welch [Baum *et al.*, 1970] permet d'estimer les paramètres $\lambda = (\pi, A, B)$ d'un HMM. Comme illustré sur la figure 5.1 :

$$\pi = P(X_0) \tag{5.9}$$
$$A = P(X_t|X_{t-1}) \tag{5.10}$$
$$B = P(Y_t|X_t) \tag{5.11}$$

Dans l'algorithme de Baum-Welch le modèle λ' est construit à partir du modèle λ de façon à maximiser la vraisemblance du modèle pour une ou plusieurs séquences d'observations $\mathcal{D} = \{\mathbf{d_l}\}$ avec $\mathbf{d_l} = (y_0, \ldots, y_T)$:

$$P(\mathcal{D}|\lambda') > P(\mathcal{D} \mid \lambda) \tag{5.12}$$

La fonction $\xi_t(i,j)$ donne la probabilité d'être dans l'état i à l'instant t et dans l'état j à l'instant $t+1$ connaissant le modèle et la séquence d'observations :

$$\xi_t^l(i,j) = P(X_t = i, X_{t+1} = j \mid \mathbf{d_l}, \lambda) \tag{5.13}$$

La probabilité $\gamma_t(i)$ d'être dans l'état i à l'instant t de la séquence d_l est donc donnée par :

$$\gamma_t^l(i) = \sum_j \xi_t^l(i,j) \tag{5.14}$$

Les formules de ré-estimation des paramètres π, A et B utilisées dans Baum-Welch sont :

$$\pi(i) = \sum_{l=1}^{L} \gamma_0^l(i) \tag{5.15}$$
$$A(i,j) = \frac{\sum_{l=1}^{L}\sum_{t=0}^{T_l} \xi_t^l(i)}{\sum_{l=1}^{L}\sum_{t=0}^{T_l} \gamma_t^l(i)} \tag{5.16}$$
$$B_i(k) = \frac{\sum_{l=1}^{L}\sum_{t=0}^{T_l} \gamma_t^l(i)\delta(y_t = k)}{\sum_{l=1}^{L}\sum_{t=0}^{T_l} \gamma_t^l(i)} \tag{5.17}$$

où $\delta(y_t = k)$ est la fonction qui vaut 1 si $y_t = k$ et 0 sinon.

5.4.2 Généralisation

Dans le cas plus général d'un DBN défini par la paire $(B_0, B_\rightarrow)$, l'apprentissage des paramètres consiste à estimer à partir d'un ensemble de séquences observées, d'un côté, les paramètres de la structure B_0 en prenant en compte le début de chaque séquence et, de l'autre, ceux de la structure $B_\rightarrow$ en s'intéressant à l'ensemble des transitions.

Expectation

Soit $\mathcal{D} = \{d_l\}_{l=1,\ldots,L}$ un ensemble de séquences observées. Chaque séquence d_l est composée d'une série de vecteurs de valeurs prises par les variables du réseau :

$$d_l = \{\mathbf{d_l^t}\}_{t=0}^{T_l}. \tag{5.18}$$

Dans un premier temps, l'étape *expectation* consiste à calculer pour chaque séquence d_l les probabilités *a posteriori* $P(Z_t^i = j, Pa(Z_t^i) = k \mid d_l)$ pour $t = 0, \ldots, T_l$.

Maximization

Paramètres de B_0 L'estimation des paramètres de B_0 maximisant la vraisemblance se fait à partir des probabilités sur les débuts de chaque séquence :

$$E\left[N_{ijk}\right] = \sum_{l=1}^{L} P(Z_0^i = j, Pa(Z_0^i) = k \mid d_l) \tag{5.19}$$

$$\theta_{ijk}^0 = \frac{\sum_{l=1}^{L} P(Z_0^i = j, Pa(Z_0^i = k \mid d_l)}{\sum_{j=1}^{q_i} \sum_{l=1}^{L} P(Z_0^i = j, Pa(Z_0^i) = k \mid d_l)} \tag{5.20}$$

Paramètres de $B_\rightarrow$ Pour $B_\rightarrow$ nous devons considérer les séquences dans leur intégralité :

$$E\left[N_{ijk}\right] = \sum_{l=1}^{L} \sum_{t=1}^{T_l} P(Z_t^i = j, Pa(Z_t^i) = k \mid d_l) \tag{5.21}$$

Ainsi les nouveaux paramètres θ_{ijk} de $B_\rightarrow$ sont donnés par :

$$\theta_{ijk}^{\rightarrow} = \frac{\sum_{l=1}^{L} \sum_{t=1}^{T_l} P(Z_t^i = j, Pa(Z_t^i) = k \mid d_l)}{\sum_{j=1}^{q_i} \sum_{l=1}^{L} \sum_{t=1}^{T_l} P(Z_t^i = j, Pa(Z_t^i) = k \mid d_l)} \tag{5.22}$$

5.5 Mélange de gaussiennes

Un mélange de gaussiennes peut être vu comme une distribution gaussienne conditionnée par une variable discrète. Ainsi dans l'exemple de la figure 5.2, la loi marginale de la variable U est un mélange de gaussiennes dont l'expression est donnée par la somme :

$$P(U = u) = \sum_k P(K = k) P(U = u | K = k) \tag{5.23}$$

$$= \sum_k \pi_k f(\mathbf{u}; \mu_\mathbf{k}, \Sigma_k) \tag{5.24}$$

où $P(K = k) = \pi_k$ et $P(U = u|K = k) = f(\mathbf{u}; \mu_\mathbf{k}, \Sigma_k)$ est une loi gaussienne de moyenne $\mu_\mathbf{k}$ et de matrice de covariance Σ_k. Nous rappelons que l'expression d'une telle gaussienne est :

$$f(\mathbf{u}, \Phi_{s,k}^j) = \frac{1}{(2\pi)^{d/2} \left|\Sigma_{s,k}^j\right|^{1/2}} \exp\left[-(\mathbf{u} - \mu_{s,k}^j)' {\Sigma_{s,k}^j}^{-1} (\mathbf{u} - \mu_{s,k}^j)\right] \tag{5.25}$$

où d est la dimension de $\mathbf{u}$ et $\left|\Sigma_{s,k}^j\right|$ est le déterminant de la matrice de covariance.

Soit $\mathcal{D} = \{\mathbf{u_i}\}_{i=1\ldots N}$ un ensemble de N échantillons d'apprentissage.

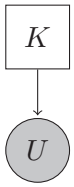


FIGURE 5.2 – Représentation graphique d'un mélange de gaussiennes.

Initialisation Un tirage aléatoire permet de définir arbitrairement une partition de la base d'apprentissage en n sous-ensembles $c_1, ..., c_n$. Il suffit par exemple de tirer pour chaque valeur de k un échantillon $\mathbf{u_k}$ comme centre du sous-ensemble c_k et d'affecter ensuite chaque échantillon de la base au sous-ensemble dont le centre est le plus proche. L'initialisation se fait ensuite en prenant :

$$\pi_k = 1 \tag{5.26}$$

$$\mu_{\mathbf{k}} = \sum_{\mathbf{u} \in c_k} \frac{\mathbf{u}}{|c_k|} \tag{5.27}$$

$$\Sigma_k = \frac{1}{|c_k|} \sum_{\mathbf{u} \in c_k} (\mathbf{u} - \mu_{\mathbf{k}})(\mathbf{u} - \mu_{\mathbf{k}})' \tag{5.28}$$

Expectation Cette étape consiste à calculer la probabilité d'appartenance de chacun des échantillons à chacune des gaussiennes. Elle est donnée par la règle de Bayes :

$$P(K = k|U = \mathbf{u}) = \frac{P(U = \mathbf{u}|K = k)P(K = k)}{\sum_c (P(U = \mathbf{u}|K = c)P(K = c))} \tag{5.29}$$

$$P(K = k|U = \mathbf{u}) = \frac{\pi_k f(\mathbf{u}; \mu_{\mathbf{k}}, \Sigma_k)}{\sum_c \pi_c f(\mathbf{u}; \mu_{\mathbf{c}}, \Sigma_c)} \tag{5.30}$$

Maximization La maximisation de la vraisemblance se fait en ré-estimant les paramètres du modèle à partir des probabilités $p_{k,\mathbf{u}} = P(K = k|U = \mathbf{u})$:

$$\pi_k = \frac{1}{N} \sum_u p_{k,\mathbf{u}} \tag{5.31}$$

$$\mu_{\mathbf{k}} = \frac{1}{\sum_{\mathbf{u}} p_{k,\mathbf{u}}} \sum_{\mathbf{u}} p_{k,\mathbf{u}} \mathbf{u} \tag{5.32}$$

$$\Sigma_k = \frac{1}{\sum_{\mathbf{u}} p_{k,\mathbf{u}}} \sum_{\mathbf{u}} p_{k,\mathbf{u}} (\mathbf{u} - \mu_{\mathbf{k}})(\mathbf{u} - \mu_{\mathbf{k}})' \tag{5.33}$$

5.6 Discussion

Nous avons rappelé dans ce chapitre le principe de l'apprentissage de paramètres par maximisation de la vraisemblance. Dans le cas où toutes les variables sont observées, l'estimation se fait directement par le calcul des fréquences d'observation des différentes valeurs prises par les variables du modèles. Dans le cas où l'observabilité est partielle, l'algorithme EM permet de réestimer itérativement les paramètres à partir de l'espérance du nombre de cas observés pour

les différentes valeurs des variables du réseau. L'étape *Expectation* consiste à estimer les valeurs prises par les variables cachées en fonction des observations et des paramètres courants du modèle. Et l'étape *Maximization* consiste à recalculer les paramètres du modèle à partir du résultat de l'étape *Expectation*. Ces deux étapes sont répétées jusqu'à la convergence de la valeur des paramètres.

Convergence vers un optimum local Il est bien connu que l'algorithme EM converge vers un maximum local de la vraisemblance du modèle. La raison est que cet algorithme se comporte comme une remontée de gradient. Le point de convergence dépend du point de départ, c'est à dire du modèle initial utilisé. La sélection du modèle initial a donc une influence sur le résultat de l'apprentissage. Il n'existe cependant pas de règle générale pour choisir la valeur des paramètres initiaux.

Sémantique des variables cachées Lorsqu'une variable est complètement cachée, comme par exemple dans le cas de la variable *Poids sec* en hémodialyse (voir chapitre 3), aucune contrainte n'est posée par les observations lors de l'apprentissage sur le sens donné aux différents états de la variable. Le problème traité est donc celui de l'apprentissage non supervisé d'une classification visant à maximiser l'homogénéité des différentes classes. Ceci nous empêche de donner *a priori* un sens à ces variables non observées. Elles peuvent donc servir de variables intermédiaires, pour augmenter la vraisemblance du modèle, mais **elles ne peuvent pas être utilisées pour réaliser un diagnostic**.

Approche fréquentiste ou bayésienne ? Comme nous l'avons évoqué dans le paragraphe 1.2, le sens donné à la mesure de probabilité dans le raisonnement bayésien n'est pas celui d'une fréquence de réalisation d'un événement mais celui d'un degré de confiance dans la vérité d'une hypothèse. Cette différence de sémantique est à l'origine d'un affrontement entre deux écoles. Chez les Bayésiens, les observations sont intégrées par l'application stricte de la règle de Bayes, alors que les Fréquentistes utilisent la loi des grands nombres pour apprendre les propriétés d'une expérience répétée plusieurs fois. Les méthodes d'apprentissage présentées ici sont fondées sur l'estimation du maximum de vraisemblance (MLE) et découlent d'une approche fréquentiste. Ainsi nous construisons de manière fréquentiste un modèle qui sera utilisé pour faire de l'inférence bayésienne. Cependant, l'approche fréquentiste n'est pas la seule envisageable. L'approche bayésienne consiste à considérer non pas la valeur des paramètres du modèle mais une distribution de probabilité sur les valeurs de ces paramètres. Elle suppose la définition d'une loi *a priori* qui est mise à jour suite à chaque observation. Les paramètres sont estimés en prenant le maximum *a posteriori* de la distribution sur les paramètres (MAP). L'utilisation d'une loi *a priori* peut être contraignante mais l'approche bayésienne présente l'avantage de pouvoir donner une estimation quand n est petit (peu de données).

Chapitre 6

Classification pour la prévention des complications de l'abord vasculaire en hémodialyse

Sommaire

L'hémodialyse est une technique de filtration extra-corporelle. L'abord vasculaire désigne le point d'accès utilisé pour réaliser la circulation extra-corporelle du volume sanguin. Il s'agit le plus souvent d'une veine du bras qui a été élargie par la création chirurgicale d'une fistule artérioveineuse. La technique consiste à établir un pont entre la veine et une artère dans le but de l'élargir et d'augmenter le débit sanguin à l'intérieur de celle-ci. Cette veine sera ensuite piquée à chaque séance de dialyse. Le nombre de veines utilisables étant limité, il est important de préserver l'abord vasculaire en traitant précocement les complications. L'une des complications les plus courantes est la sténose de l'abord vasculaire, c'est à dire un rétrécissement qui peut se produire à différents endroits de l'abord et se manifeste par une modifications des contraintes sur la circulation sanguine autour des points d'accès.

Nous présentons dans ce chapitre un modèle dynamique paramétré par apprentissage supervisé permettant de classifier des enregistrements de séances de dialyses selon le risque de présence d'une sténose, c'est à dire d'un rétrécissement, au niveau de l'abord vasculaire. Ce travail [11] est le résultat d'une collaboration entre la société Diatélic, l'INRIA et la société Gambro.

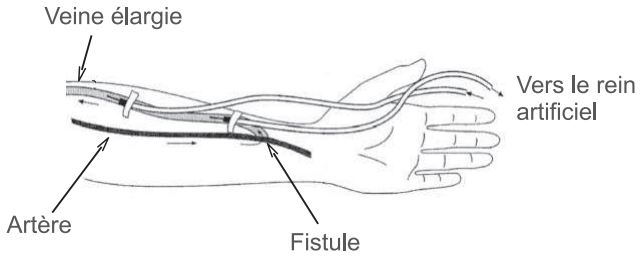


FIGURE 6.1 – Représentation schématique de l'accès vasculaire utilisé pour la dialyse extracorporelle.

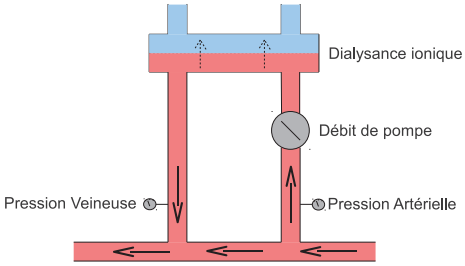


FIGURE 6.2 – Circuit de l'épuration extracorporelle.

6.1 Problématique

La circulation extracorporelle mise en place lors d'une séance d'hémodialyse est un circuit sanguin artificiel connecté à la circulation naturelle par un accès vasculaire comme représenté sur la figure 6.1.

Comme illustré par la figure 6.2, la circulation extracorporelle se fait à partir de l'accès vasculaire grâce à une pompe qui envoie le sang du patient vers le dialyseur où les échanges se font avec un liquide de dialyse appelé dialysat par l'intermédiaire d'une membrane séparant les deux compartiments.

La dynamique du fluide au niveau de la circulation extracorporelle dépend bien évidement du réglage de la pompe, mais elle dépend également de la forme de l'accès vasculaire, de la façon dont celui-ci a été piqué, du type d'aiguille utilisé, de la viscosité du sang, ... Une évolution de la forme de l'accès vasculaire, en particulier un rétrécissement, aura donc des conséquences visibles sur la dynamique de la circulation. Une évolution du rapport entre le débit et les pressions artérielle et veineuse pourra donc être l'indicateur d'une sténose à un certain endroit de l'accès vasculaire. En plus d'avoir des conséquences sur la dynamique, des modifications de la circulations peuvent avoir un impact sur la qualité de la filtration, en particulier en cas de phénomènes de recirculation comme illustré sur la figure 6.3. La dialysance ionique mesure un débit d'épuration de certains ions exprimé en ml/min. C'est la quantité de sang totalement épurée de ces ions chaque minute. Il a été montré que cette mesure était un bon indicateur de la clairance de l'urée qui représente le volume de sang totalement épuré de l'urée chaque minute. La dialysance ionique est utilisée comme indicateur de la qualité de la filtration.

L'objectif de la surveillance de l'accès vasculaire est de détecter suffisamment tôt un rétrécissement éventuel de l'accès pour qu'une intervention puisse avoir lieu avant que celui-ci soit

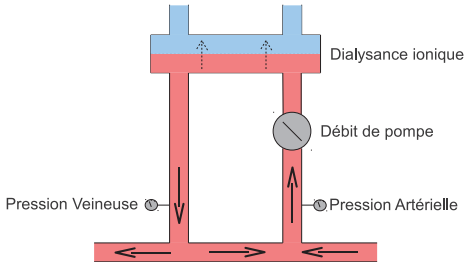


FIGURE 6.3 – Le phénomène de recirculation peut être l'indicateur d'un rétrécissement en aval de l'accès vasculaire. Le volume sanguin réellement épuré chaque minute est alors diminué et ne correspond pas au volume mesuré au niveau de la pompe.

définitivement perdu. Les mesures de pressions, de débit et de dialysance sont réalisées par la machine de dialyse au cours de chaque séance. Ces mesures sont interprétables par un expert humain et leur évolution permet d'estimer l'état de l'accès vasculaire. Il est ainsi possible de distinguer trois grandes classes de données caractérisées par un score allant de 1 à 3 :

- le score 0 est donné dans le cas d'un patient stable dont les mesures sont normales sur une période suffisamment longue ;
- le score 1 est donné dans le cas d'un patient présentant un ou plusieurs paramètres anormaux mais stables, c'est à dire sans aggravation récente ;
- le score 2 est donné aux patients nécessitant de façon urgente des examens complémentaires et éventuellement une intervention médicale.

Les machines de dialyses fournissent, à chaque séance et pour chaque patient, des données qui, si elles sont analysées, permettent de diagnostiquer précocement une aggravation de l'état de l'abord vasculaire. Cependant la surveillance en continu nécessite d'analyser chaque jour un grand nombre de données. Le but de ce travail est d'automatiser la classification des données afin de permettre au médecin de ne regarder en priorité que les cas les plus urgents.

6.2 Approche

Ce travail a été réalisé en collaboration avec Bernard BENE, expert de cette problématique et auteur d'un brevet [7] Gambro sur le suivi de l'accès vasculaire en hémodialyse. Nous disposions pour construire le système d'un ensemble de règles formalisant le processus de décision suivi par l'expert. Ces règles définissent un ensemble de critères permettant l'attribution du score 0 et du score 2. Les données ne répondant ni aux critères de stabilité et de normalité du score 0, ni aux critères d'évolutions défavorables du score 2 sont classées comme score 1. Comme nous l'illustrerons dans la section 6.5, ces règles pourtant bien définies ne permettent pas de construire un système fiable en raison de la grande variabilité des situations.

Nous disposons également d'une base de données étiquetées par un expert humain, composée d'un ensemble de séquences d'observations $\mathcal{D}^i$ chacune associée à un score x_i caractérisant l'état de l'abord vasculaire en fin de séquence. Chaque séquence $\mathcal{D}^i$ correspond à un enregistrement d'une durée pouvant aller de trois semaines à six mois des séances de dialyse d'un même patient. A chaque séance D_t^i sont enregistrés :

- la pression artérielle PA_t^i, qui est la pression (négative) en entrée de la pompe ;

7. Brevet No FR2911417 publié le 18/07/2008

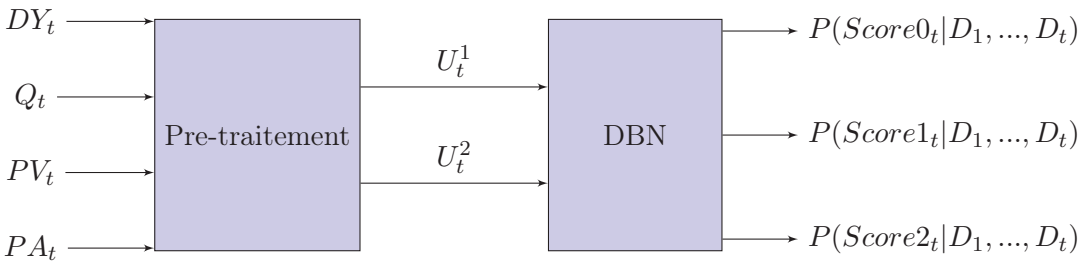


FIGURE 6.4 – Le processus d'analyse des données repose sur une première étape de pré-traitement faisant ressortir l'information pertinente sous la forme d'*evidences* utilisées ensuite dans le modèle bayésien.

- la pression veineuse PV_t^i, qui est la pression (positive) au retour vers la circulation naturelle ;
- le débit sanguin Q_t^i dans la circulation extra corporelle ;
- la dialysance ionique DY_t^i.

Il ressort de la discussion avec l'expert que les paramètres ne sont pas indépendants. La fourchette de normalité des pressions dépend en effet du débit sanguin. Plus le débit est grand, plus nous nous attendons à observer des pressions importantes. La dialysance ionique est, elle aussi, liée au débit. Plus le débit est important, plus la dialysance attendue est élevée. Par ailleurs, il apparaît dans la définition des scores que l'interprétation se fait à la fois sur les valeurs mesurées et sur leur variations.

Nous proposons donc de réaliser l'analyse automatique en deux étapes comme illustré sur la figure 6.4. La première consiste en un pré-traitement des données visant à faire ressortir une information pertinente. Dans une seconde étape, la probabilité de chaque score connaissant une séquence d'observations pourra être évaluée à l'aide d'un modèle stochastique.

6.3 Pré-traitement

Le pré-traitement consiste à vérifier la validité des données et à calculer, d'une part, un vecteur U_1 résultant de la concaténation des données brutes et, d'autre part, un vecteur U_2 reflétant l'évolution des paramètres par rapport à une référence calculée itérativement. Le vecteur U_1 est de dimension quatre. Ses dimensions reflètent :

- le débit sanguin,
- la dialysance,
- la pression artérielle,
- la pression veineuse.

Le vecteur U_2 est de dimension trois. Il est composé à partir des variations de :

- dialysances,
- pressions artérielles,
- pressions veineuses,

conditionnées par le débit sanguin.

La figure 6.5 représente en détails le processus de pré-traitement. Ce processus est une succession de filtres. Le filtre de validation *Valid* consiste à vérifier que la mesure appartient à un domaine de validité pré-défini pour chaque paramètre. Si ce n'est pas le cas, une information

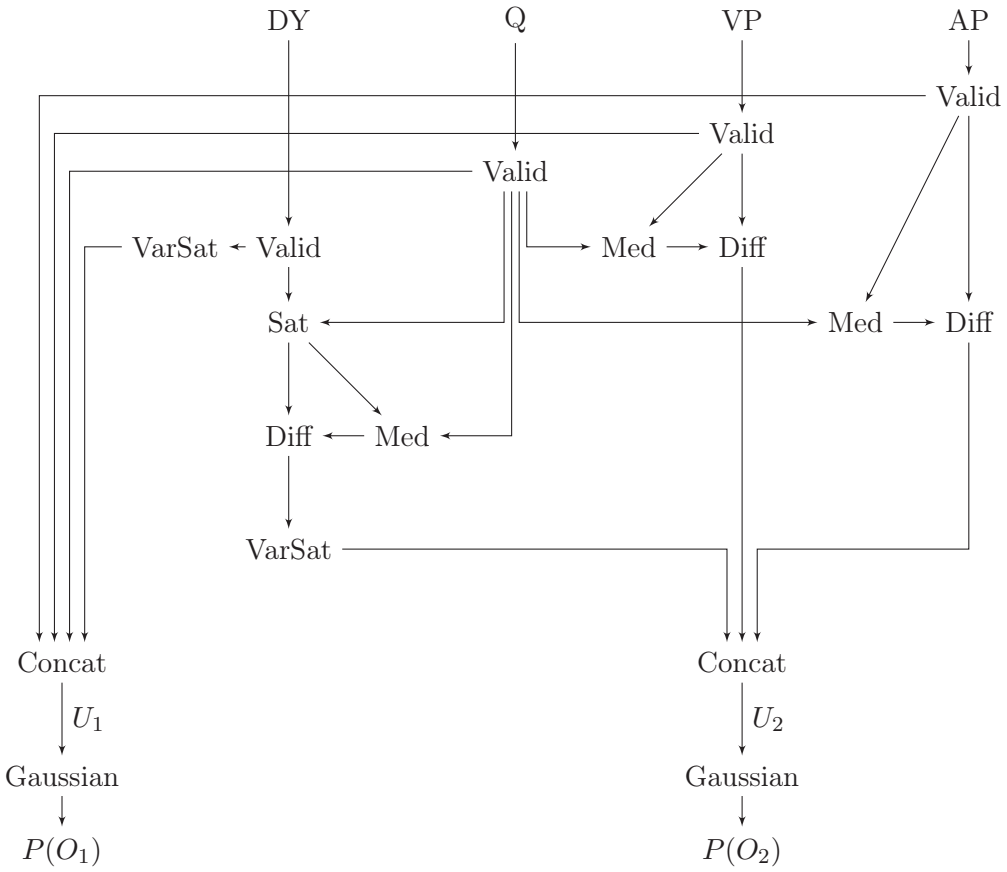


FIGURE 6.5 – Description fonctionnelle du pré-traitement des données.

donnée manquante sera propagée. Le filtre *Med* calcule une référence pour la dialysance et les pressions. Cette valeur de référence est conditionnée par la valeur de débit sanguin. Cette référence est calculée comme la médiane des valeurs observées sur les 6 derniers mois pour chaque tranche de débit sanguin. Le filtre *Diff* permet de mesurer l'écart entre la valeur courante et la référence calculée d'après les valeurs passées. Le filtre *VarSat* est utilisé pour borner par le haut ou par le bas les valeurs dans un intervalle calculé en fonction de la variance des valeurs observées. Ce filtre sert à faciliter la modélisation des données par des distributions gaussiennes et à diminuer l'impact des mesures extrêmes. L'interprétation de la mesure de dialysance n'est en effet pas symétrique. Une dialysance « trop bonne » ne doit pas être considérée comme un signe de score 2. La saturation des valeurs de dialysance très élevées nous permet d'éviter des erreurs de classification liée à la modélisation par des distributions gaussiennes comme illustré par la figure 6.6.

6.4 Modèle

Nous utilisons une structure dynamique simple (voir figure 6.7) formant une chaîne de Markov cachée. Les observations U_t^1 et U_t^2 en sortie de la phase de pré-traitement sont deux variables continues suivant des distributions apprises de manière supervisée. O_t^1 et O_t^2 sont des variables intermédiaires discrètes permettant de « découpler » légèrement les lois de U_t^1 et U_t^2 vis à

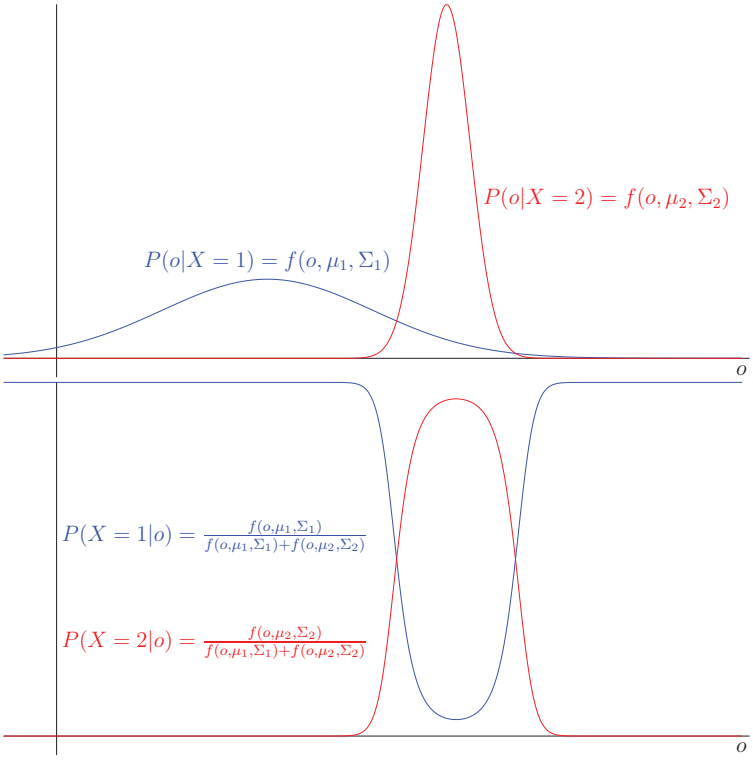


FIGURE 6.6 – Le premier graphe représente deux lois gaussiennes correspondant à deux états différents d'une variable X. La loi de X *a posteriori* est représentée en dessous. Cette figure illustre le phénomène d'inversion de l'interprétation lorsque la loi de variance la plus grande dépasse, ici à droite, la loi de plus petite variance.

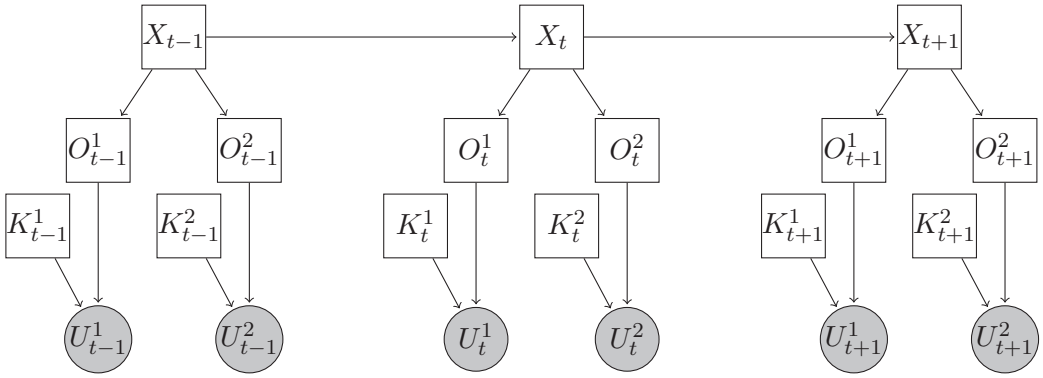


FIGURE 6.7 – DBN pour l'estimation de l'état de l'abord vasculaire.

vis du modèle dynamique qui lui est défini manuellement. La variable X_t représente le score correspondant à la séance de dialyse D_t.

6.4.1 Mélange de gaussiennes

La fonction d'observation utilisée est un mélange de gaussiennes multivariées (voir paragraphe 5.5). Dans notre application, chaque score s est caractérisé par deux mélanges de gaussiennes. Le premier est de dimension quatre et représente la variable U^1. Le second est de dimension trois et représente la variable U^2 :

$$P(U^j = \mathbf{u}|O^j = s) = \sum_{k=1}^{n} \pi^j_{s,k} f(\mathbf{u}, \Phi^j_{s,k}) \tag{6.1}$$

où $f(\mathbf{u}, \Phi^j_{s,k})$ est la gaussienne multivariée de paramètres $\Phi^j_{s,k} = (\mu^j_{s,k}, \Sigma^j_{s,k})$, $\mu^j_{s,k}$ étant le vecteur des moyennes et $\Sigma^j_{s,k}$ la matrice de covariances.

Le score x_i attribué à chaque séquence D^i par l'expert porte sur la fin de la séquence. Il est ainsi possible de définir une base d'exemples pour apprendre les distributions gaussiennes correspondant à chaque score à partir des derniers échantillons de chaque séquence. Cette notion de « fin de séquence » n'est cependant pas parfaitement bien définie. Le score porte le plus souvent sur la dernière séance, parfois sur l'avant-dernière et plus rarement sur l'avant-avant-dernière. Nous prenons donc en compte dans la base d'apprentissage les trois derniers échantillons de chaque séquence en donnant un peu plus d'importance aux séances les plus récentes en comptant le dernier échantillon trois fois, l'avant-dernier deux fois et l'avant-avant-dernier une fois.

Pour chaque score s et chaque variable U^j, l'apprentissage du mélange de gaussiennes se fait avec un algorithme de type EM maximisant la vraisemblance du modèle. Les paramètres à apprendre sont ($\pi^j_{s,k}$, $\Sigma^j_{s,k}$) et $\mu^j_{s,k}$ pour chaque valeur de $k = 1...n$, n étant le nombre de gaussiennes défini arbitrairement. Nous avons empiriquement établi la complexité du modèle à 3 gaussiennes ($n = 3$), en maximisant la qualité de la reconnaissance sur notre base de travail.

La base de données d'apprentissage dont nous disposons contient 1458 séquences avec la répartition suivante :

- 1003 cas de score 0 (68.8%)
- 388 cas de score 1 (26.6%)

– 67 cas de score 2 (4.6%)

Comme souvent dans les applications médicales, les cas pathologiques sont, fort heureusement, relativement rares. Ce déséquilibre important entre le nombre de cas normaux et le nombre de cas pathologiques rend la tâche de modélisation plus délicate pour deux raisons. La première est qu'il est difficile, voire impossible, d'observer l'ensemble des situations pathologiques. La seconde est qu'il y a un *a priori* fortement déséquilibré entre les deux classes, qui aurai tendance, dans un cas limite, à nous faire opter pour la classe score 0 alors que, du point de vu de l'application, marquer comme score 0 un cas de score 2 est l'erreur la plus grave.

6.4.2 Modèle dynamique

Les mélanges de gaussiennes sont ensuite insérés dans le modèle dynamique. Nous nous trouvons dans le cas particulier évoqué dans la section 1.5.1, car les nœuds continus sont uniquement au niveau des feuilles du réseau. L'inférence peut donc se faire en considérant un ensemble d'*evidences* incertaines pour les variables O_t^1 et O_t^2 calculées à partir des lois continues.

Les tableaux 6.1 et 6.2 décrivent les lois conditionnelles associant les variables O_t^1 et O_t^2 à la variable d'état X_t. La loi dynamique $P(X_{t+1}|X_t)$ est décrite dans le tableau 6.3. La loi dynamique que nous utilisons correspond simplement à une hypothèse de continuité du processus. Les variables O_t^1 et O_t^2 ont été introduites pour pouvoir ajuster le poids relatif des fonctions d'observations (apprises de manière statique) vis à vis de la loi dynamique.

	$O_t^1 = 0$	$O_t^1 = 1$	$O_t^1 = 2$
$X_t = 0$	0.95	0.05	0.0
$X_t = 1$	0.025	0.95	0.025
$X_t = 2$	0.0	0.05	0.95

TABLE 6.1 – $P(O_t^1|X_t)$

	$O_t^2 = 0$	$O_t^2 = 1$	$O_t^2 = 2$
$X_t = 0$	0.95	0.05	0.0
$X_t = 1$	0.025	0.95	0.025
$X_t = 2$	0.0	0.05	0.95

TABLE 6.2 – $P(O_t^2|X_t)$

	$X_{t+1} = 0$	$X_{t+1} = 1$	$X_{t+1} = 2$
$X_t = 0$	0.80	0.10	0.10
$X_t = 1$	0.10	0.80	0.10
$X_t = 2$	0.10	0.10	0.80

TABLE 6.3 – $P(X_{t+1}|X_t)$

6.5 Résultats

La base de données étiquetée par l'expert humain a été divisée en deux par l'expert lui même. La première partie étant utilisée pour l'apprentissage des gaussiennes et la seconde pour

Auto \Humain	Score 0	Score 1	Score 2
Score 0	763	55	1
Score 1	80	196	15
Score 2	8	20	66

TABLE 6.4 – Comparaison des scores résultants de l'analyse automatique avec ceux de l'analyse manuelle.

	Sensibilité	VPP
Score 0	0.89	0.93
Score 1	0.72	0.67
Score 2	0.80	0.70

TABLE 6.5 – Evaluation des résultats du système d'analyse automatique en termes de sensibilité et de VPP.

l'évaluation. Nous confrontons pour l'évaluation des résultats le score donné en fin de séquence par l'analyse automatique avec le score donné par l'expert humain. La base de données de test est composée de 1204 enregistrements avec la répartition suivante :

- 851 cas de score 0 (70.5%)
- 271 cas de score 1 (22.5%)
- 82 cas de score 2 (7%)

Le tableau 6.4 récapitule les différences entre les sorties de l'analyse automatique et les résultats attendus par l'expert humain. Dans ce tableau les éléments diagonaux correspondent aux enregistrements pour lesquels le score donné par le modèle informatique est en accord avec celui de l'expert humain.

A partir du tableau de contingence, nous pouvons mesurer la sensibilité et la valeur prédictive positive (VPP) de la détection de chaque score en considérant les scores donnés par l'expert humain comme référence. Le calcul est le suivant :

$$\text{Sensibilité}(Score_i) = \frac{\text{Vrais } Score_i \text{ détectés}}{\text{Nombre total de vrais } Score_i} \tag{6.2}$$

$$\text{VPP}(Score_i) = \frac{\text{Vrais } Score_i \text{ détectés}}{\text{Nombre total de score } Score_i \text{ détectés}} \tag{6.3}$$

Les résultats en termes de sensibilités et de VPP sont donnés dans le tableau 6.5.

A titre de comparaison, nous avons utilisé les règles de décision telles qu'elles ont été formalisées par l'expert humain dans une première description du problème. Les résultats de l'évaluation du système à base de règles sont donnés dans les tableaux 6.6 et 6.7. La grande variabilité des signaux observés fait que l'utilisation d'une telle approche est inadaptée. Il est bien sûr difficile de formaliser le raisonnement opéré par l'humain avec un système de règles. Les données en entrée du système sont des valeurs numériques alors que le raisonnement humain est essentiellement symbolique. Ainsi la notion de stabilité d'un paramètre, perçue visuellement par l'humain, est difficile à mettre en équation. Par ailleurs, le savoir faire de l'expert repose sur beaucoup de connaissances non dites, acquises par l'expérience. Il est donc d'une certaine façon plus pertinent de travailler par mimétisme, en modélisant la fonction de décision à partir des résultats de l'analyse humaine, plutôt que de demander à l'humain de formaliser son propre raisonnement.

Règles \Humain	Score 0	Score 1	Score 2
Score 0	724	135	25
Score 1	118	120	42
Score 2	9	16	15

TABLE 6.6 – Comparaison des scores obtenus d'un coté avec le système à base de règles et de l'autre par l'expert humain.

	Sensibilité	VPP
Score 0	0.85	0.82
Score 1	0.44	0.43
Score 2	0.18	0.38

TABLE 6.7 – Sensibilité et VPP de la détection de chaque score avec le système à base de règles.

6.6 Conclusion

Nous avons présenté dans ce chapitre une application des techniques d'apprentissage à un problème de diagnostic médical. La contribution est applicative puisque le modèle ainsi développé a été jugé suffisamment pertinent pour pouvoir être distribué par Gambro dans les centres de dialyse, comme module intégré à une solution logicielle plus globale. Le modèle développé permettra ainsi de tirer parti d'informations issues des machines de dialyse qui ne sont pas du tout prises en compte aujourd'hui, à la fois parce que leur analyse nécessite une formation particulière et parce que la quantité de données générée est trop importante pour qu'un personnel médical puisse s'intéresser chaque jour l'ensemble des patients suivi. L'apport de ce système est de pouvoir proposer une sélection de patients triés par ordre de gravité.

Sur le plan méthodologique nous avons pu ainsi constater la pertinence d'aborder par apprentissage un problème de diagnostic lorsque l'on dispose de données étiquetées, même partiellement (ici, seules les fin de séquences étaient étiquetées). L'intérêt de l'apprentissage est qu'il permet de prendre en compte des connaissances non dites acquises par l'humain. Remarquons cependant que la validation est réalisée uniquement en comparaison avec l'analyse humaine. La validité des diagnostics n'est bien sûr pas démontrée vis à vis de l'état réel du patient.

Chapitre 7

Apprentissage de modèles pour la segmentation d'électrocardiogrammes

Sommaire

Le travail de segmentation, qui consiste à positionner des marqueurs sur un signal temporel est souvent nécessaire en médecine pour localiser dans le temps des phénomènes et mesurer leur durée. Ce problème peut concerner toutes sortes de signaux de différente nature comme par exemple les mouvements respiratoires, l'activité électrique cardiaque ou neuronale, ... Selon l'objectif de la segmentation, la nature éventuellement variable du signal, l'analyse peut nécessiter une expertise particulière. Nous nous intéressons ici à la possibilité d'apprendre par l'exemple les critères de positionnement des marqueurs.

Ce chapitre présente une approche par apprentissage pour réaliser le découpage d'un électrocardiogramme (ECG) selon les différentes phases du cycle cardiaque. Nous montrons ainsi que l'approche stochastique permet de mimer le comportement d'un expert humain en reproduisant sa façon d'analyser un signal dynamique. Ce travail [1] est le fruit d'une collaboration entre l'INRIA et la société Cardiabase dont le cœur de métier est l'analyse de signaux ECG produits dans le cadre d'études pharmaceutiques.

dI	mesure bipolaire entre le bras droit et le bras gauche.
dII	mesure bipolaire entre le bras droit et la jambe gauche.
$dIII$	mesure bipolaire entre le bras gauche et la jambe gauche.
aVR	mesure unipolaire sur le bras droit.
aVL	mesure unipolaire sur le bras gauche.
aVF	mesure unipolaire sur la jambe gauche.
$V1$	mesure sur le $4^{\text{ème}}$ espace intercostal droit, sur le bord droit du sternum.
$V2$	mesure sur le $4^{\text{ème}}$ espace intercostal gauche, sur le bord gauche du sternum.
$V3$	mesure entre $V2$ et $V4$.
$V4$	mesure sur le $5^{\text{ème}}$ espace intercostal gauche, sur la ligne médioclaviculaire.
$V5$	mesure sur le $4^{\text{ème}}$ espace intercostal gauche, sur la ligne axillaire antérieure.
$V6$	mesure sur le $4^{\text{ème}}$ espace intercostal gauche, sur la ligne axillaire moyenne.

TABLE 7.1 – Les 12 dérivations d'un ECG standard

7.1 Contexte

Dans le cadre du développement de nouveaux médicaments les laboratoires pharmaceutiques sont tenus de réaliser un certain nombre d'études et d'essais cliniques en vue d'obtenir l'autorisation de mise sur le marché (AMM). Les essais sur l'homme se décomposent en plusieurs phases dont la première consiste à vérifier les niveaux de tolérance et de toxicité de la nouvelle molécule. Ces essais de *phase I* sont réalisés sur un petit nombre de sujets sains (de 20 à 200) recevant des doses croissantes du médicament sous surveillance médicale. Cette surveillance porte notamment sur le système cardiovasculaire afin de détecter d'éventuels effets indésirables de la molécule.

L'électrocardiogramme est un enregistrement de l'activité électrique à l'origine des mouvements cardiaques. Il permet de mettre en évidence les différentes phases du cycle cardiaque qui est composé des systoles (contractions auriculaire et ventriculaire) et de la diastole (pendant laquelle le muscle se décontracte pour collecter le sang). La surveillance porte en particulier sur la durée des différentes phases extraites à partir de l'ECG.

L'enregistrement est réalisé à l'aide d'un certain nombre d'électrodes en contact avec la peau au niveau du thorax, des bras et de la jambe gauche. A partir de ces électrodes sont enregistrés un certain nombre de signaux unipolaires (différence de potentiel entre une électrode et un point de référence) et bipolaires (différence de potentiel entre deux électrodes). Un ECG standard est ainsi constitué de 12 signaux comme détaillés dans le tableau 7.1. Un exemple de tracé des 12 pistes sur une période est donné par la figure 7.1.

La figure 7.2 représente les différentes ondes et les différents intervalles caractéristique de l'ECG mettant en évidence les différentes phases de l'activité cardiaque. L'onde P correspond à la phase de dépolarisation des oreillettes (systole auriculaire). Elle est suivie par le complexe QRS provoquant la contraction des ventricules (systole ventriculaire). Finalement l'onde T correspond à la repolarisation des ventricules (diastole). La durée totale de l'activité ventriculaire est mesurée par l'intervalle QT qui représente l'indicateur le plus utilisé en pharmacologie clinique.

Les essais cliniques réalisés par les laboratoires pharmaceutiques sont ainsi générateurs d'une très grande quantité d'enregistrements dont la lecture repose sur la mesure des différents intervalles. La société Cardiabase, spécialisée dans la lecture et l'analyse d'enregistrements ECG, traite chaque année de nombreux enregistrements en s'appuyant sur une procédure semi-automatique.

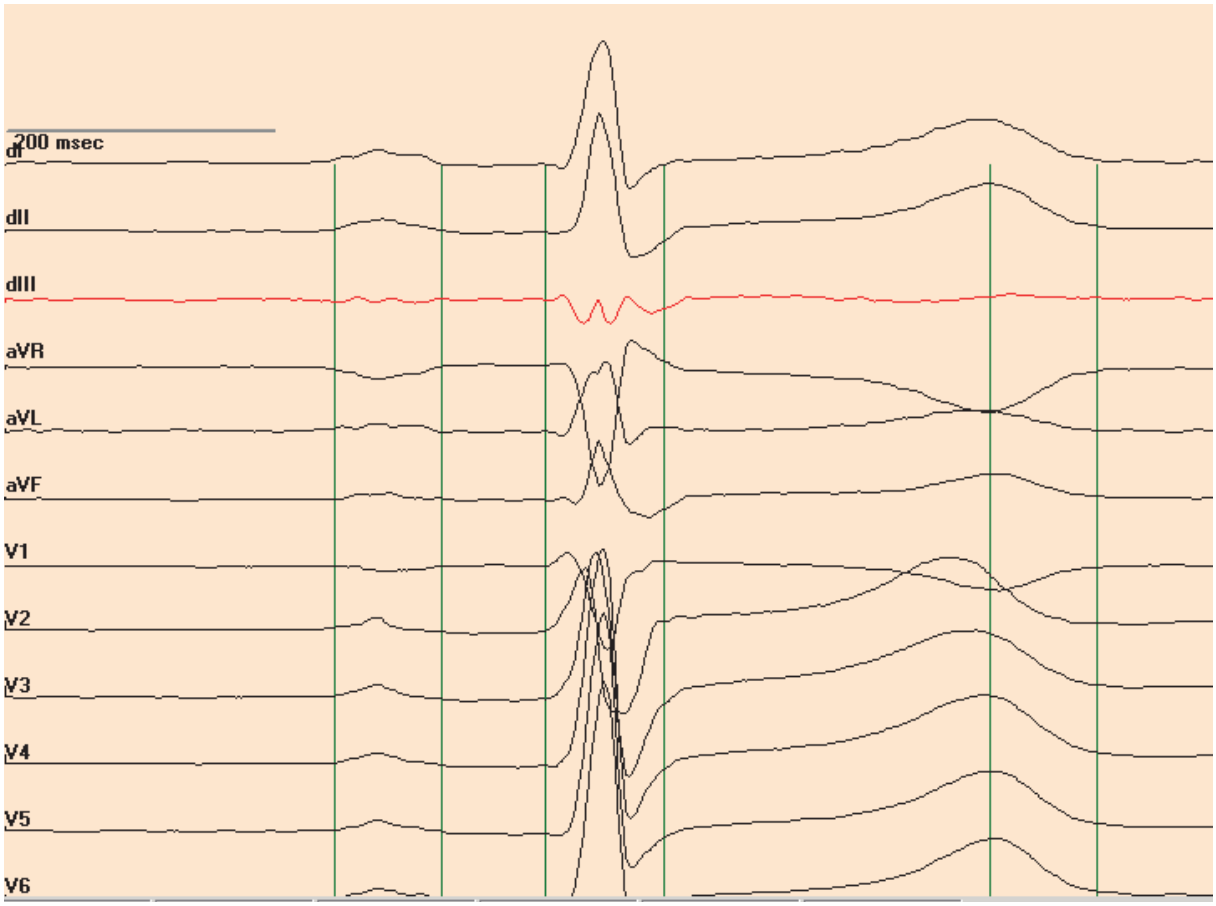


FIGURE 7.1 – Exemple de tracé des 12 pistes sur une période.

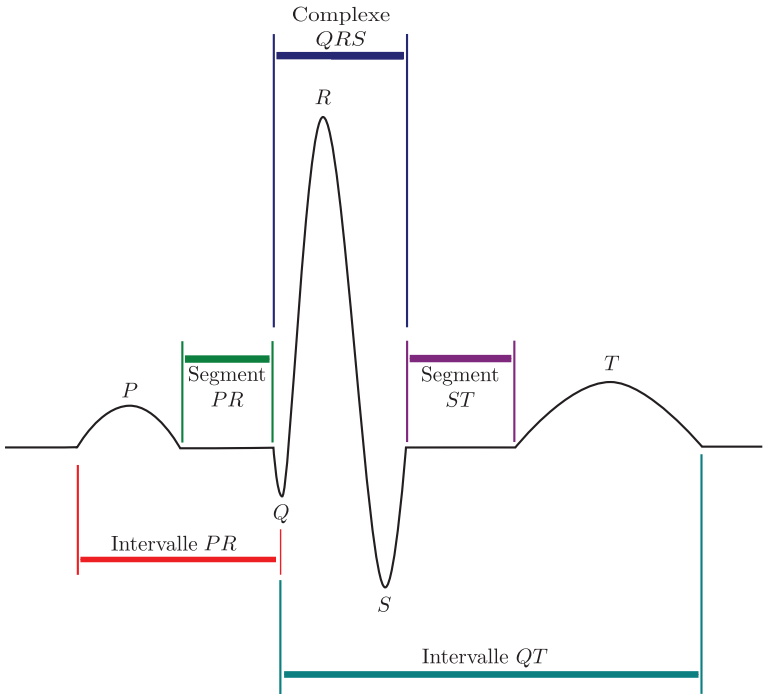


FIGURE 7.2 – Les différents segments de l'ECG

Avant d'être validés par des cardiologues, les enregistrements sont pré-traités par un logiciel puis revus par des opérateurs formés à la lecture d'ECG. Cette révision permet de corriger les défauts de la segmentation automatique.

Dans la pratique, les enregistrements ECG sont très variables d'un sujet à l'autre ce qui rend l'automatisation de la segmentation difficile. Les techniques classiques de segmentation automatique ont en effet tendance à reproduire systématiquement les mêmes erreurs sur un sujet donné. Les enregistrements sont suffisamment spécifiques de la personne pour être considérés par certains comme une empreinte personnelle, au même titre que l'empreinte digitale. Une personne habituée à voir les enregistrements d'un sujet peut en effet parfois reconnaitre le sujet d'après les spécificités de ses enregistrements.

En s'inspirant des techniques de reconnaissance de la parole, nous proposons de construire par apprentissage un modèle stochastique du signal ECG afin de pouvoir adapter celui-ci à un ou plusieurs sujets et être capable d'apprendre une segmentation plus pertinente, évitant ainsi les corrections manuelles répétitives. Nous proposons de modéliser les enregistrements ECG à l'aide d'un ensemble de chaînes de Markov cachées représentant les différents segments du signal. Nous proposons conjointement une méthode de classification non supervisée des enregistrements permettant de regrouper les enregistrements similaires et de développer une approche multi-modèles.

7.2 Travaux connexes

Dans le domaine de l'analyse automatique de l'ECG, il semble pertinent de distinguer le problème de la détection du complexe QRS, dont une revue a été proposé par Köhler et al

[Köhler *et al.*, 2002], du problème plus difficile de la segmentation. Bien que donnant souvent de bon résultats pour la détection du complexe QRS, les méthodes par seuillage manquent de souplesse pour le problème de la segmentation. Il faut par ailleurs citer le problème de la classification d'ECG qui consiste le plus souvent à distinguer les enregistrements pathologiques des enregistrements normaux. Ce problème est parfois abordé en utilisant des chaînes de Markov cachées [Kičmerová *et al.*, 2005].

7.2.1 Utilisation d'une représentation temps-fréquences

Différents travaux montrent l'intérêt d'utiliser une représentation en temps-fréquences des signaux ECG pour effectuer la segmentation [Martinez *et al.*, 2004] [Khawaja *et al.*, 2005]. Il est en effet souvent plus pertinent de travailler en observant les variations du signal plutôt que le signal lui même. Dans le cas de la segmentation de l'ECG ce sont principalement les discontinuités de l'enregistrement qui permettent de déterminer le début et la fin des différentes ondes. Ces discontinuités se traduisent par l'existence de hautes fréquences dans le signal.

La transformée de Fourier permet de déterminer la puissance d'un signal périodique dans les différents domaines de fréquence. Une technique souvent utilisée pour avoir une représentation locale en fréquences du signal, consiste à effectuer la transformée sur une fenêtre glissante. Bien entendu, la longueur de la fenêtre considérée a un impact à la fois sur la précision de la localisation et sur le domaine de fréquences visible. En effet des fenêtres plus grandes permettent de mettre en évidence des fréquences plus basses alors que les fenêtres courtes donnent une information mieux localisée dans le temps. La transformée de Fourier à fenêtre glissante du signal $x(t)$ est donnée par l'équation 7.1.

$$TFFG(\omega, \tau) = \int_{-\infty}^{+\infty} x(t) g(t-\tau) \exp(j\omega t) dt \tag{7.1}$$

où $g(t)$ est une fonction de fenêtrage, non nulle pour une petite période de temps autour de 0. ω désigne la fréquence de la transformation et τ est un facteur de translation.

A l'instar de la transformée de Fourier, la transformée en ondelettes consiste à décomposer le signal en une somme pondérée de signaux élémentaires. La différence principale avec la transformée de Fourier est que, dans le cas de la transformée en ondelettes, ces signaux élémentaires sont localisés dans le temps. La transformée en ondelettes peut ainsi être vue comme une généralisation de la transformée de Fourier à fenêtre glissante, dans laquelle la fenêtre est de longueur variable et est incluse dans le signal élémentaire. Le signal élémentaire appelé ondelette mère peut avoir différentes formes (sinusoïdale comme dans le cas de l'ondelette de Morlet, carrée comme dans le cas de l'ondelette de Haar, ...) et subit une dilatation, remplaçant la notion de fréquence de la transformée de Fourier par la notion d'échelle. L'expression de la transformée en ondelette continue du signal $x(t)$ est donnée par :

$$\gamma(a, b) = \frac{1}{\sqrt{|a|}} \int_{-\infty}^{+\infty} x(t) \overline{\psi(\frac{t-b}{a})} dt \tag{7.2}$$

où $\psi(t)$ est l'ondelette mère, exprimée en représentation complexe et $\overline{\psi(t)}$ désigne son conjugué. L'ondelette de Haar, représentée figure 7.3 est définie par :

$$\psi(x) = \begin{cases} 1 & \text{si } 0 \leq x < \frac{1}{2} \\ -1 & \text{si } \frac{1}{2} \leq x < 1 \\ 0 & \text{sinon} \end{cases} \tag{7.3}$$

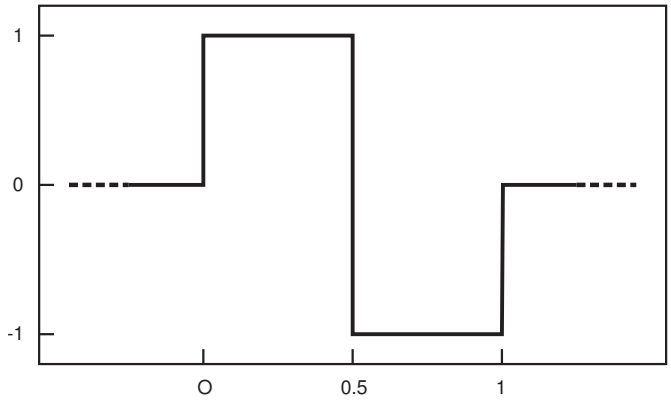


FIGURE 7.3 – Ondelette mère de Haar.

Sa forme carrée est particulièrement intéressante dans le cadre de la segmentation d'électrocardiogrammes car les hautes fréquences qu'elle intègre permettent de mettre en évidence les singularités du signal. La figure 7.4 représente la transformée en ondelette de Haar de la piste dII.

7.2.2 Utilisation de modèles stochastiques

Le problème de la segmentation de l'ECG peut être abordé en utilisant des modèles de Markov cachés. L'approche à base de HMM n'est cependant pas triviale comme le montre le travail de Hughes et al. [Hughes *et al.*, 2004]. L'approche proposée utilise un état par segment de l'ECG et les auteurs montrent qu'il est nécessaire d'ajouter une représentation explicite de la durée des états pour améliorer la qualité de la segmentation. La représentation utilisant un état unique par segment du signal est probablement insuffisante pour avoir une représentation markovienne du signal.

Graja et Boucher ont obtenu de bon résultats en utilisant une transformée en ondelettes du signal associée à des arbres de Markov cachés [Graja and Boucher, 2003]. Cette approche multi-échelle prend en compte dans un premier temps la dépendance hiérarchique des coefficients avant d'intégrer la notion de dépendance temporelle entre coefficients successifs d'une même échelle.

7.3 Utilisation de plusieurs chaînes de Markov inter-connectées

Dans notre application nous partons de signaux ECG déjà découpés en périodes chacune décomposables en 6 segments (voir figure 7.5) :

- la ligne de base $Base1$,
- l'onde P,
- la ligne de base $Base2$,
- le complexe QRS,
- l'onde T,
- la ligne de base $Base3$.

Plutôt que d'associer un état à chaque segment, nous proposons de représenter chaque partie du signal par un HMM complet. Le modèle Λ du signal complet est alors constitué de la concaténa-

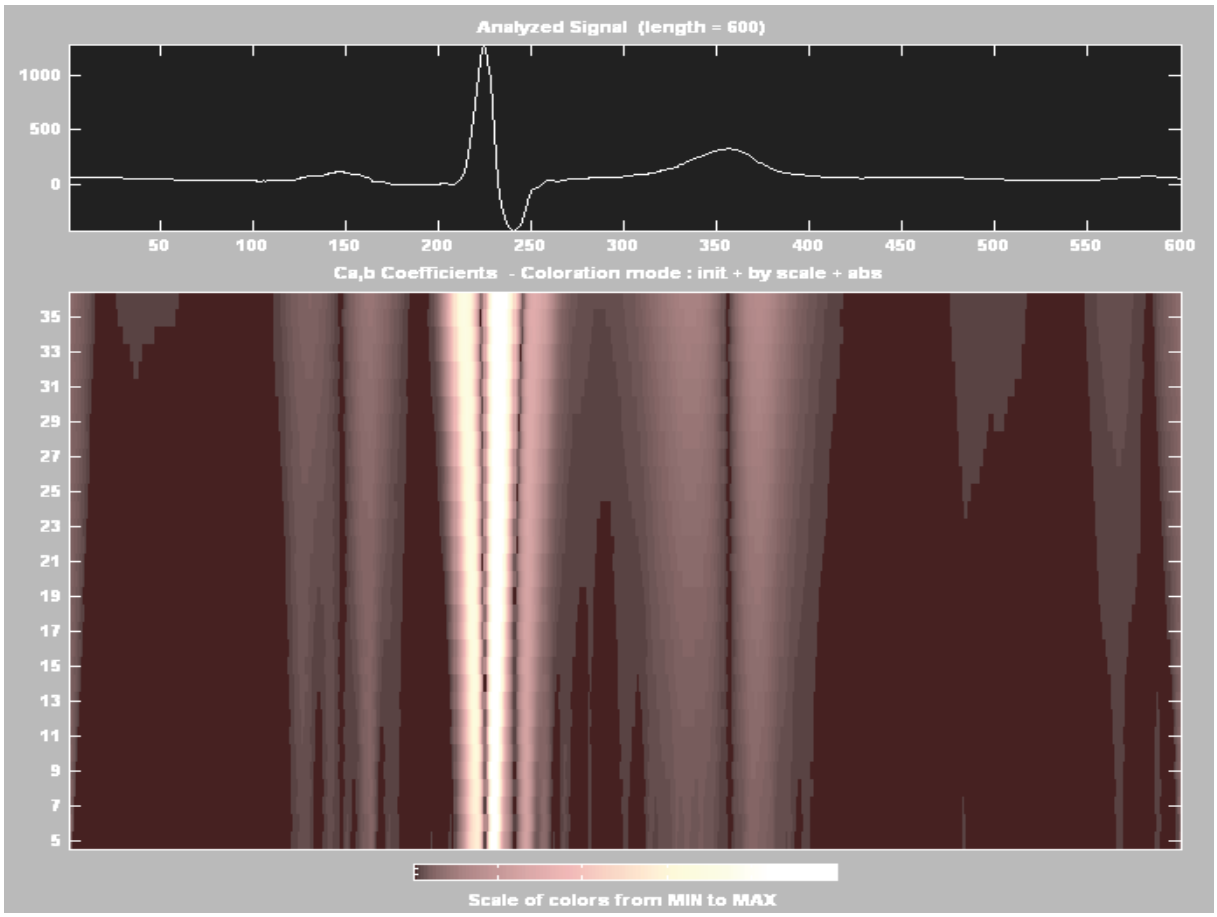


FIGURE 7.4 – Transformée en ondelette de Haar de la piste dII.

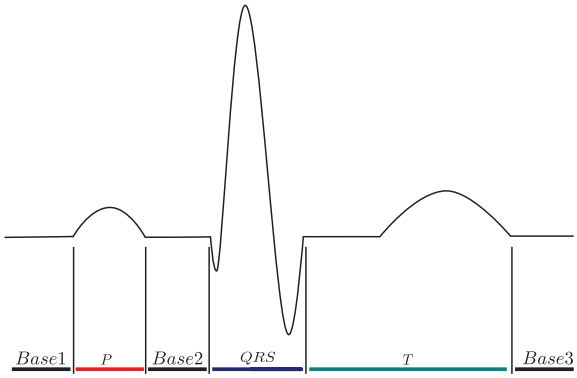


FIGURE 7.5 – Les 6 sous-segments de la base d'apprentissage.

tion de 6 sous-modèles λ_{Base1}, λ_P, λ_{Base2}, λ_{QRS},λ_T et λ_{Base3}. Chaque sous-modèle est construit par apprentissage sur une base de signaux déjà segmentés. Nous avons choisi d'imposer une structure de transitions gauche-droite à 5 états (voir figure 7.6). Remarquons que les 5 états de chaque modèle correspondent à 5 valeurs d'un processus stochastique discret. Ainsi, les représentations graphiques données dans ce chapitre ne sont pas des DBN mais des automates représentant la matrice de transition. L'absence d'arc dans le graphe traduit une probabilité nulle de transition entre les deux états. Les modèles sont à densité gaussienne appris à l'aide d'un algorithme EM (voir chapitre 5). Si une transition particulière a une probabilité nulle dans le modèle initial, celle-ci ne peut être observée lors de l'étape *Expectation* et la probabilité de cette transition restera nulle après l'étape de *Maximization*. Ainsi la maximisation de la vraisemblance par l'algorithme EM ne modifie pas la structure des transitions.

Le modèle global peut être considéré comme un HMM hiérarchique [Fine and Singer, 1998] dont le premier niveau serait constitué de 6 états abstraits, un pour chaque segment, et dans lequel chacun de ces macro-états serait associé au sous-modèle λ_i correspondant. Murphy a montré que les HMM hiérarchiques (HHMM) trouvaient une représentation naturelle sous la forme de DBNs en introduisant une variable par niveau hiérarchique [Murphy, 2002]. Dans le modèle global Λ nous nous intéressons aux transitions entre les différents sous-modèles. Etant donné la relative simplicité de notre HHMM il nous parait plus simple de le présenter comme la concaténation des 6 sous-modèles. Cela revient à construire un nouveau processus stochastique qui comporte alors 30 états (voir figure 7.7). Les probabilités de transition entre deux sous-modèles consécutifs sont choisies égales aux probabilités de sortie du modèle précédent.

7.4 Segmentation sur 12 pistes.

Comme nous l'avons mentionné précédemment, l'ECG est composé de 12 pistes enregistrées simultanément. L'approche utilisée par l'expert humain consiste à superposer les 12 enregistrements comme illustré sur la figure 7.8 et à positionner les marqueurs en fonction de l'enveloppe des 12 signaux. Plusieurs solutions sont envisageables pour intégrer l'information provenant des 12 signaux dans la segmentation. Si l'on considère que les observations sont conditionnellement indépendantes connaissant l'état de la chaîne, nous pouvons insérer plusieurs variables observées dans le modèle, une par piste, chacune étant conditionnée par l'état de la chaîne. C'est une façon naturelle d'aborder le problème. L'hypothèse d'indépendance conditionnelle est cependant peu

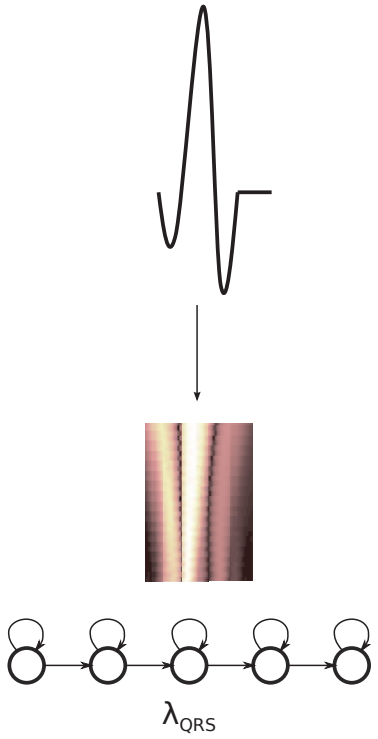


FIGURE 7.6 – Un HMM est construit pour chaque segment. Les coefficients issus de la transformée en ondelettes servent à apprendre les paramètres du modèle correspondant au segment (ici le complexe QRS).

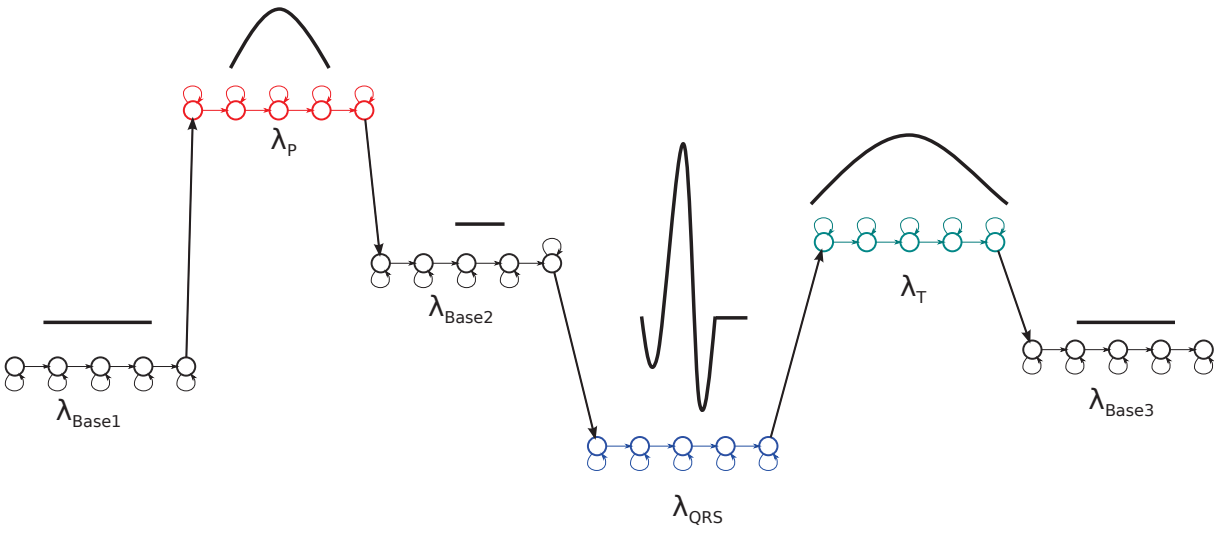


FIGURE 7.7 – Le modèle global Γ est construit en concaténant les 6 sous-modèles λ_i.

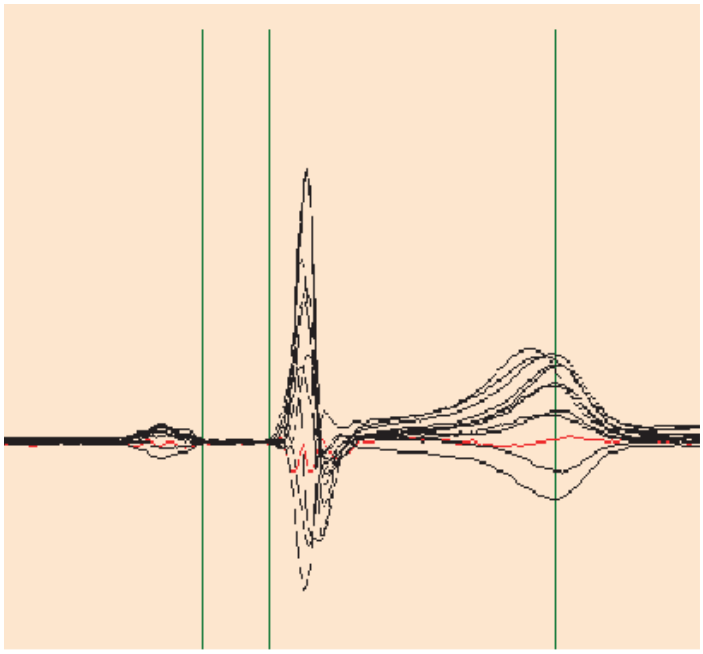


FIGURE 7.8 – Superposition des 12 pistes de l'enregistrement, réalisée lors de la segmentation manuelle de l'ECG.

vraisemblable étant donné que, dans le processus d'enregistrement de l'ECG, certaines électrodes sont communes à plusieurs pistes.

Une autre approche envisageable consisterait à augmenter la dimension de la variable observée en concaténant les coefficients provenant des transformées en ondelettes des 12 signaux. Cette approche présente l'inconvénient d'être lourde numériquement puisque la matrice de covariance de la gaussienne multivariée grossit comme le carré de la dimension du vecteur observé. De plus, l'intégration dans une même variable de dimensions fortement corrélée peut s'avérer problématique numériquement pour l'inversion de la matrice de covariance. Ce problème est connu et peut se traiter par exemple en réduisant le nombre de dimensions à l'aide d'une analyse en composantes principales.

Nous avons finalement obtenu de bon résultats en traitant les 12 pistes séparément. Le traitement séparé des 12 pistes nous amène à 12 propositions pour chaque marqueur que nous filtrons ensuite en prenant la valeur médiane. Le fait d'utiliser la médiane nous évite de tenir compte de résultats éventuellement aberrants.

7.5 Approche multi-modèles

Comme nous l'avons mentionné précédemment les enregistrements ECG sont fortement variables d'une personne à l'autre. Ainsi nous pouvons constater une amélioration des résultats en apprenant un modèle spécifique pour chaque sujet. Dans la pratique, nous souhaitons pouvoir obtenir une segmentation de qualité sur un sujet non encore rencontré dans la base d'apprentissage. L'idée que nous envisageons ici est de former des classes de sujets et de construire un modèle par classe.

Nous n'avons cependant pas d'information *a priori* sur la ressemblance des enregistrements dans notre base d'apprentissage. Un des algorithmes les plus connus pour la classification non supervisée de donnée est l'algorithme des *k-means* qui est une méthode de partitionnement de l'espace en un nombre donné de classes. Le problème peut également être abordé de manière hiérarchique en agglomérant successivement les données similaires. Une revue des méthodes de classification non supervisée a été proposée par Jain et al [Jain *et al.*, 1999].

Etant donné une séquence observée et une collection de modèles dynamiques, la vraisemblance de chaque modèle est souvent utilisée pour décider à quel modèle correspond le mieux la séquence. Ainsi, disposant d'une collection de modèles $\{\Lambda_i\}_{i=1}^{N}$ pour segmenter un ECG particulier s, nous choisirons le modèle ayant la plus grande vraisemblance $\arg\max_{\Lambda_i} P(s \mid \Lambda_i)$ pour traiter le signal. Nous proposons donc d'utiliser également la vraisemblance comme critère de constitution des classes d'ECG. L'algorithme que nous proposons d'utiliser est une adaptation naturelle de l'algorithme des *K-Means* où les classes sont représentées par un HMM et où la notion de distance est remplacée par celle de vraisemblance (voir algorithme 4). Cet algorithme, parfois appelé *K-models*, consiste à partir d'une partition arbitraire des enregistrements à apprendre le HMM représentatif de chaque groupe d'enregistrements et à réaffecter itérativement les enregistrements à la classe dont le HMM a la plus grande vraisemblance.

L'idée d'utiliser un HMM comme représentant d'un ou plusieurs signaux temporels pour aborder le problème de la classification a été évoquée par Juang et Rabiner qui ont proposé une mesure de distance entre modèles [Juang and Rabiner, 1985]. La question de la distance entre HMM a également été abordée par Falkhausen et al [Falkhausen *et al.*, 1995]. Smyth s'est intéressé à la question de l'initialisation de l'algorithme et à la recherche du nombre de classes [Smyth, 1997]. Li et Biswas ont présenté une vue globale du problème intégrant la recherche du nombre de classes et, pour chaque classe, la recherche du nombre d'états dans le HMM maximisant un critère bayésien de sélection des modèles [Li and Biswas, 2000]. Concernant les applications Krogh et al ont proposé l'utilisation de HMM pour traiter le problème de la modélisation de protéines [Krogh *et al.*, 1994]. Citons également le travail de Butler qui s'est intéressé à la classification non supervisée de signaux acoustiques à l'aide de HMM [Butler, 2003].

7.6 Résultats

Nous présentons dans ce paragraphe une comparaison des segmentations obtenues sur 173 enregistrements ECG par différentes méthodes :

Cardio - la segmentation manuelle validée par plusieurs cardiologues servant de référence.

Ad-hoc - un algorithme spécifique utilisé précédemment dans le procédé de la société Cardiabase.

Generic - la segmentation par HMM utilisant un modèle unique.

Cluster - la segmentation obtenue par l'approche multi-modèles en utilisant 10 classes.

L'apprentissage des modèles a été réalisé sur une base de donnée validée de 1800 enregistrements. La comparaison statistique est faite sur la durée estimée par les différentes méthodes de l'intervalle PR, du complexe QRS et de l'intervalle QT.

Les tableaux 7.2, 7.4 et 7.6 représentent la dispersion des durées des différents intervalles selon la méthode d'estimation. Les figures 7.9, 7.10 et 7.11 donnent une représentation graphique des dispersions.

Entrées : Un ensemble d'enregistrements $\mathcal{S} = \{s_i\}_{i=1}^{N}$, le nombre K de classes
`/* Sélectionner` K `enregistrements distincts pour l'initialisation. */`
pour $j = 1$ *à* K **faire**
  $C_j = \{s_{ij}\}$;
fin
tant que *les classes* $\{C_j\}$ *ne sont pas stables* **faire**
  pour $j = 1$ *à* K **faire**
    Apprendre $\Lambda_j = EM(C_j)$;
    pour $i = 1$ *à* N **faire**
      Calculer $L_{i|j} = \log P(s_i \mid \Lambda_j)$;
    fin
    $C_j \leftarrow \emptyset$;
  fin
  pour $i = 1$ *à* N **faire**
    $j^* = \arg\max_j L_{i|j}$;
    $C_{j^*} = C_{j^*} \bigcup \{s_i\}$;
  fin
  $iter \leftarrow iter + 1$;
fin

Algorithme 4: Classification non supervisée de signaux temporels en fonction de la vraisemblance du HMM représentatif de la classe.

Intervalle *PR*. La méthode *Ad hoc* aboutit à des longueurs moyenne et médiane de l'intervalle *PR* proches de celles du cardiologue. Cependant la dispersion des mesures est beaucoup plus importante que dans la distribution de référence. La méthode *Generic* donne des durées surestimées de 11 ms en moyenne, avec une forte dispersion des valeurs. A l'inverse, la méthode *Cluster* donne une distribution tout a fait comparable au niveau de la moyenne, de la médiane et de l'écart-type. Cette analyse se confirme par l'étude des échantillons appariés présentée dans le tableau 7.3. L'écart-type de l'erreur est en effet de 5.8 ms (3.6% de la longueur moyenne de l'intervalle) pour la méthode *Cluster* contre 28 ms (17,6 %) pour la méthode *Ad hoc* et 27.3 ms (17,1 %) pour la méthode *Generic*.

PR	Moyenne	Min.	Max.	Médiane	Ecart type
Ad hoc	160.6	110	408	158	33.3
Cardio	159.2	114	204	158	19.0
Generic	170.1	119	370	166	31.7
Cluster	161.9	120	210	161	18.6

TABLE 7.2 – Analyse statistique de l'intervalle *PR*.

Complexe *QRS*. Les trois méthodes sous-estiment la longueur de l'intervalle par rapport à l'estimation faite par le cardiologue (voir figure 7.10). Et la comparaison 2 à 2 des méthodes (voir tableau 7.5) ne révèle pas de différences flagrantes entre les trois approches. L'écart-type de l'erreur est d'environ 7 ms (7 % de la longueur de l'intervalle).

Δ PR	Moyenne	Ecart type	Min.	Max.
Cardio-Ad hoc	-1.3	28.0	-254	36
Cardio-Generic	-10.9	27.3	-233	18
Cardio-Cluster	-2.7	5.8	-26	21
Cluster-Ad hoc	1.3	28.3	-258	32
Generic-Ad hoc	9.5	39.3	-255	233
Cluster-Generic	8.2	25.8	-7	219

TABLE 7.3 – Comparaison 2 à 2 des méthodes de segmentation de l'intervalle PR

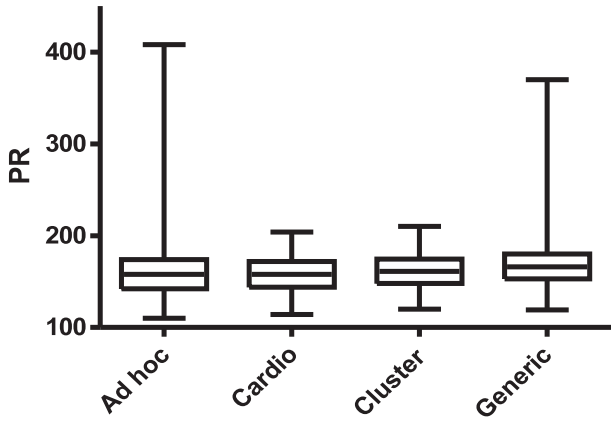


FIGURE 7.9 – Représentation des dispersions pour l'intervalle PR.

QRS	Moyenne	Min.	Max.	Médiane	Ecart type
Ad hoc	84.2	64	122	84	9.5
Cardio	95.5	76	116	96	8.3
Generic	89.3	76	107	89	6.1
Cluster	88.3	69	109	87	6.9

TABLE 7.4 – Analyse statistique du complexe QRS.

Δ QRS	Moyenne	Ecart type	Min.	Max.
Cardio-Ad hoc	11.3	7.2	-16	32
Cardio-Generic	6.1	6.4	-15	26
Cardio-Cluster	7.1	6.9	-12	30
Cluster-Ad hoc	4.1	8.0	-25	24
Generic-Ad hoc	5.2	7.2	-21	25
Cluster-Generic	1.0	4.2	-20	10

TABLE 7.5 – Comparaison 2 à 2 des méthodes de segmentation du complexe QRS.

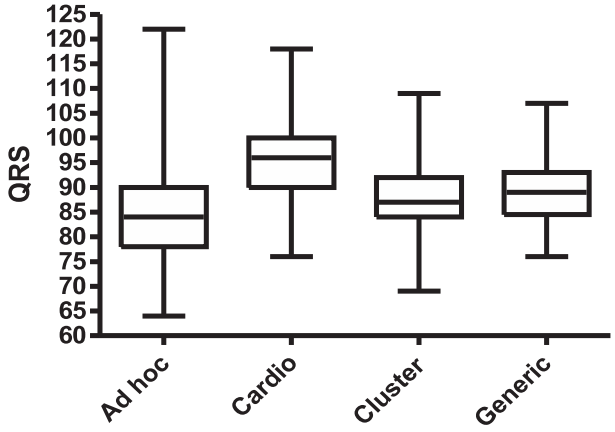


FIGURE 7.10 – Représentation des dispersions pour le complexe QRS.

Intervalle QT. Sur l'intervalle QT, les 3 méthodes de segmentation donnent des résultats similaires avec un petit avantage pour la méthode *Cluster*. Celle-ci donne en effet une estimation de la longueur moyenne légèrement meilleure que la méthode *Ad hoc* et un écart-type des différences plus petit qu'avec la méthode *Generic* (voir tableau 7.7).

QT	Moyenne	Min.	Max.	Médiane	Ecart type
Ad hoc	396.5	306	454	398	25.8
Cardio	400.5	330	474	402	25.6
Generic	403.1	339	483	404	26.3
Cluster	399.0	330	481	400	25.3

TABLE 7.6 – Analyse statistique de l'intervalle QT.

Δ QT	Moyenne	Ecart type	Min.	Max.
Cardio-Ad hoc	4.0	7.0	-18	24
Cardio-Generic	-2.5	16.3	-142	36
Cardio-Cluster	1.5	8.8	-30	27
Cluster-Ad hoc	2.5	9.6	-29	34
Generic-Ad hoc	6.5	17.9	-54	166
Cluster-Generic	4.1	14.2	-25	142

TABLE 7.7 – Comparaison 2 à 2 des méthodes de segmentation de l'intervalle QT.

7.7 Conclusion

Nous avons développé une approche par apprentissage permettant d'améliorer la segmentation des ECGs en comparaison avec la méthode *ad hoc* pré-existante. L'apprentissage est rendu possible par l'existence d'exemples de signaux segmentés par des cardiologues. L'approche multi-modèles permet une mesure plus fine des intervalles PR et QT. Les différences de résultats entre

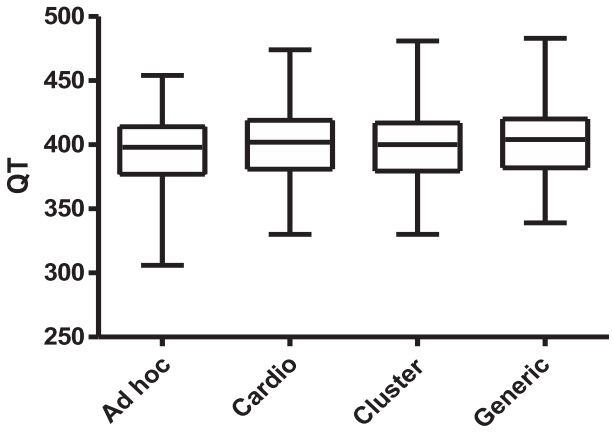


FIGURE 7.11 – Représentation des dispersions pour l'intervalle QT.

les différentes approches sont moins flagrantes sur l'intervalle QRS. Le fait que cette approche multi-modèles donne de meilleurs résultats nous montre l'intérêt de développer des modèles plus spécifiques d'un sous-groupe de sujets.

La constitution des sous-groupes de sujets est entièrement réalisée de manière non supervisée et se fonde sur les similarités existant entre les différents enregistrements. Le prétraitement réalisé sur le signal n'est pas spécifique aux ECGs. Ainsi, l'approche proposée présente l'intérêt d'être complètement générique et de reposer entièrement sur les exemples utilisés pour l'apprentissage sans qu'aucune connaissance au préalable sur la forme des signaux ne soit utilisée. La segmentation automatique fonctionne par mimétisme, en essayant d'obtenir un résultat proche de celui de l'humain. Ceci rend l'approche transposable à d'autres problématiques comme par exemple l'analyse de signaux respiratoires ou électroencéphalographiques.

Chapitre 8

Apprentissage par renforcement à partir d'exemples

Sommaire

Un certain nombre de problèmes médicaux peuvent être considérés comme des problèmes de contrôle dans lesquels le but est de définir une stratégie de traitement pour amener le patient dans un certain état. En reprenant l'exemple de l'hémodialyse, la prescription du poids sec est une action sur l'état d'hydratation du patient que le néphrologue cherche à réguler. Nous

pouvons également citer le problème du dosage de l'érythropoïétine (EPO) administrée aux patients insuffisants rénaux en compensation de la déficience due à l'insuffisance rénale. L'EPO est impliquée dans la fabrication des globules rouge et l'adaptation de la prescription a pour but de réguler le taux de globules rouges dans le sang du patient.

Les processus de décision markovien (MDPs) permettent de formaliser des problèmes de contrôle en considérant un agent en interaction avec un environnement (le système à contrôler). L'agent agit sur l'environnement et perçoit une récompense qui dépend à la fois de son action et de l'état du système. Cette récompense caractérise le but à atteindre. L'apprentissage par renforcement consiste pour l'agent à apprendre, par l'expérience, une politique, c'est à dire une stratégie d'actions, maximisant la récompense perçue à long terme.

Nous nous intéressons dans ce chapitre au problème de l'apprentissage par renforcement dans le cadre particulièrement contraint des applications médicales. Le problème que nous posons ici est d'apprendre à contrôler un système, que nous assimilons au patient, qui a pour caractéristiques d'être **partiellement observable** et surtout **partiellement explorable**. Cette problématique rejoint de manière plus générale celle de l'utilisation pratique des techniques d'apprentissage par renforcement dans un cadre réel.

8.1 Processus de décision markoviens

Cette section a pour objectif de rappeler la définition des principaux éléments utilisés dans le cadre des MDPs auxquels nous ferons référence dans la suite du chapitre.

8.1.1 Définition générale

De manière générale, un MDP est défini formellement par :
- le processus aléatoire S_t prenant ses valeurs dans l'ensemble $\mathcal{S}$ des états perçus par l'agent,
- le processus aléatoire R_t associé à la récompense prenant ses valeurs dans $\mathbb{R}$,
- le processus décisionnel A_t prenant ses valeurs dans l'ensemble $\mathcal{A}$ des actions possibles,
- une fonction de transition $\mathcal{T}$ définie sur l'espace des transitions $\mathcal{S} \times \mathcal{A} \times \mathcal{S}$ par

$$\mathcal{T}^a_{ss'} = P(S_{t+1} = s' | S_t = s, A_t = a), \tag{8.1}$$

- l'espérance de récompense $\mathcal{R}$ définie également sur l'espace des transitions $\mathcal{S} \times \mathcal{A} \times \mathcal{S}$ par

$$\mathcal{R}^a_{ss'} = E(R_{t+1} | S_t = s, A_t = a, S_{t+1} = s). \tag{8.2}$$

La représentation graphique de cette définition est donnée par la figure 8.1.

8.1.2 Politique

On appelle politique π_t la mesure de probabilité définie à chaque pas de temps sur $\mathcal{S} \times \mathcal{A}$ par

$$\pi_t(s, a) = P(A_t = a | S_t = s). \tag{8.3}$$

Elle donne la probabilité avec laquelle l'agent choisit l'action a dans l'état s à l'instant t. Cette définition est celle d'une politique non déterministe.

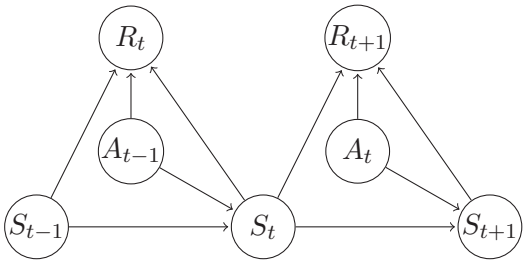


FIGURE 8.1 – Représentation graphique d'un MDP

Remarquons qu'elle permet également de représenter des politiques déterministes en prenant

$$\pi_t(s, a) = \delta_{a_t}(a) \tag{8.4}$$

où δ_{a_t} est la distribution Dirac centrée en a_t et $a_t = \pi_t(s)$ est l'action choisie selon une politique déterministe.

La définition d'une politique se traduit graphiquement par l'ajout d'un arc entre S_t et A_t.

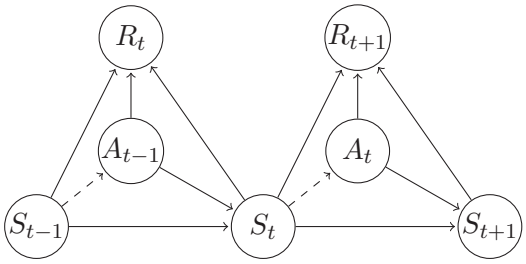


FIGURE 8.2 – Représentation graphique d'un MDP avec une politique fixée

8.1.3 Valeur d'état

La notion de valeur d'état est au centre des algorithmes d'apprentissage dans les MDPs. La valeur d'un état se mesure comme l'espérance de récompense à partir de cet état en suivant une politique π. On appelle *fonction de valeur* la fonction définie sur l'espace d'états $\mathcal{S}$ pour une politique π par :

$$V^{\pi}(s) = E_{\pi}\left[\sum_{k=0}^{\infty} \gamma^k R_{t+1+k} | S_t = s\right]. \tag{8.5}$$

Elle permet de définir une relation d'ordre partielle sur les politiques :

$$\pi \leq \pi' \text{ ssi } V^{\pi}(s) \leq V^{\pi'}(s) \text{ pour tout } s \in \mathcal{S}. \tag{8.6}$$

Il existe une ou plusieurs politiques notées π^* qui dominent toutes les autres et qui partagent la fonction de valeur optimale :

$$V^*(s) = \max_{\pi} V^{\pi}(s) \text{ pour tout } s \in \mathcal{S}. \tag{8.7}$$

8.1.4 Valeur d'action

La notion de valeur d'action est similaire à la notion de valeur d'état. Etant donné une politique π la valeur $Q^\pi(s,a)$ est l'espérance de la récompense reçue par l'agent effectuant l'action a dans l'état s et suivant ensuite la politique π. On appelle *Q-fonction* la fonction définie sur $\mathcal{S} \times \mathcal{A}$ par :

$$Q^\pi(s,a) = E_\pi \left[\sum_{k=0}^{\infty} \gamma^k R_{t+1+k} \middle| S_t = s, A_0 = a \right]. \tag{8.8}$$

Connaissant la valeur des actions, il est possible de déduire la valeur des états d'après la politique π :

$$V^\pi(s,a) = \sum_a \pi(s,a) Q^\pi(s,a). \tag{8.9}$$

La Q-fonction optimale est définie par :

$$Q^*(s,a) = \max_\pi Q^\pi(s,a) \text{ pour tout } s \in \mathcal{S} \text{ et } a \in \mathcal{A}. \tag{8.10}$$

Remarquons que la valeur d'une politique optimale en s est la valeur de la meilleure action en s :

$$V^*(s) = \max_a Q^*(s,a). \tag{8.11}$$

Connaissant la Q-fonction optimale, il est donc facile de déduire une politique optimale π^* en recherchant pour chaque état les actions maximisant la valeur. Nous utilisons ici la notation $\pi(s)$ pour désigner l'action a telle que $\pi(s,a) = 1$.

$$\pi^*(s) \in argmax_a Q^*(s,a). \tag{8.12}$$

8.1.5 Opérateur de Bellman

Il est possible de décomposer la valeur d'un état a (8.8) comme la somme de l'espérance de récompense immédiate ($k = 0$) et de l'espérance de récompense future ($k = 1...\infty$) qui peut s'exprimer en fonction de la valeur des états s' résultant de la transition depuis s.

$$Q^\pi(s,a) = E_\pi \left[R_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^k R_{t+2+k} \middle| S_t = s, A_t = a \right] \tag{8.13}$$

$$Q^\pi(s,a) = \sum_{s'} \mathcal{T}^a_{ss'} E_\pi \left[R_{t+1} + \gamma \sum_{k=0}^{\infty} \gamma^k R_{t+2+k} \middle| S_t = s, S_{t+1} = s', A_t = a \right] \tag{8.14}$$

$$Q^\pi(s,a) = \sum_{s'} \mathcal{T}^a_{ss'} \left[\mathcal{R}^a_{ss'} + \gamma V^\pi(s') \right] \tag{8.15}$$

De (8.15) et (8.11) nous pouvons déduire l'équation de Bellman :

$$V^*(s) = \max_a \sum_{s'} \mathcal{T}^a_{ss'} \left[\mathcal{R}^a_{ss'} + \gamma V^*(s') \right] \tag{8.16}$$

Il est possible d'écrire de manière similaire pour la Q-fonction optimale :

$$Q^*(s,a) = \sum_{s'} \mathcal{T}^a_{ss'} \left[\mathcal{R}^a_{ss'} + \gamma \max_{a'} Q^*(s',a') \right] \tag{8.17}$$

L'équation de Bellman permet d'exprimer V^* comme le point fixe d'un opérateur contractant. V^* est alors l'unique solution de cette équation.

8.2 Le problème posé

Le contexte applicatif visé par ce travail nous impose des contraintes fortes sur l'observation de l'état du patient et sur l'exploration des actions thérapeutiques. La prise de décision est supposée réalisée à temps discret. Les actions sont supposées discrètes. L'adaptation de la dose d'un traitement est en effet souvent considérée de manière discrète sous la forme d'actions symboliques du type « augmenter ou diminuer la dose de 25% ». Les mesures réalisées sont supposée être de nature continue, ou éventuellement hybrides (association de variables observées continues et discrètes). L'objectif que nous nous fixons est de rechercher une politique maximisant une fonction de récompense, donnée *a priori* et calculée d'après les observations.

8.2.1 Observabilité Partielle

L'observabilité partielle est un problème récurent dans les systèmes réels. En robotique, en plus du problème de l'incertitude inhérente à la mesure, les capteurs embarqués ne donnent bien souvent qu'une information locale sur l'environnement. Il faut donc faire face à des problèmes d'*aliasing* c'est à dire des perceptions identiques provenant de configurations différentes de l'environnement. Certains paramètres peuvent ne pas être accessibles à la mesure. C'est le cas par exemple du volume d'eau corporelle dans le problème de l'hémodialyse.

Au delà des problèmes de mesure, la notion même d'état semble délicate à appréhender s'agissant du monde réel. Dans le cadre des MDPs, ou bien même des POMDPs, l'espace d'état fait partie de la définition du problème et le terme *modèle* désigne habituellement la fonction de transition qui peut ne pas être connue. Dans le cas d'un système réel, le choix de l'espace d'état du MDP fait partie de la modélisation et le terme *modèle* prend alors un sens plus large, désignant à la fois le choix des grandeurs mesurées et l'espace d'état qui sert de support à la fonction de transition et à la politique. L'environnement réel est continu, ou au moins perçu comme tel aux échelles des systèmes que nous cherchons à contrôler et surtout ses dimensions rendent impossible toute représentation complète de son état. Ceci nous amène à choisir une représentation nécessairement simplifiée du système. Nous pouvons donc différencier l'état véritable X du système de la représentation S utilisée par l'agent. La figure 8.3 illustre ce processus de modélisation dans lequel l'état S du MDP est une projection d'un espace beaucoup plus grand et où les transitions $(S_t \rightarrow S_{t+1})$ sont en fait des projections de transitions $(X_t \rightarrow X_{t+1})$ du système réel. La représentation du système réel par un modèle simplifié engendre de l'indéterminisme quant à son évolution dû à la perte d'information causée par la modélisation.

8.2.2 Exploration Partielle

Dans le problème que nous posons, la dynamique du système n'est pas supposée connue. Si des connaissances d'experts existent souvent sur le système à contrôler, elles ne sont, en général, pas suffisantes pour pouvoir être exprimées précisément dans l'espace d'état du modèle. Par ailleurs, il n'est souvent pas possible de définir un modèle universel. Si l'on s'intéresse au corps humain, tous les individus ne réagissent pas de la même façon à un même traitement sans qu'il

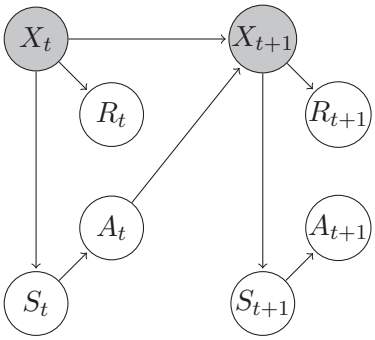


FIGURE 8.3 – Distinction entre le vrai état du système X et la représentation utilisée par l'agent S

soit toujours possible d'en connaitre les raisons. C'est pourquoi nous cherchons à apprendre à contrôler le système sans connaissance *a priori* de sa dynamique.

L'apprentissage de la politique devra donc se baser sur une exploration du système, mais qui restera partielle du fait des contraintes liées au temps imparti et surtout à la volonté d'explorer le système tout en maintenant son intégrité. Ces conditions prennent un sens évident dans le cas où le système est un être humain qu'il faut soigner en minimisant les risques, mais il s'agit certainement de conditions rencontrées de manière générale dans les systèmes réels. Si l'on cherche à apprendre à contrôler un véhicule, certaines actions ne pourront pas être explorées sans risquer l'accident.

Nous faisons donc l'hypothèse que nous disposons d'un ensemble d'exemples de contrôle. Il s'agira de l'historique des actions thérapeutiques d'un ou de plusieurs patients dans le cas d'applications médicales. Ces exemples précèdent l'apprentissage et nous supposons qu'il n'est pas possible de compléter l'exploration pendant l'apprentissage. Nous nous rapprochons donc des problèmes d'apprentissage dits « hors ligne » dans le sens où l'apprentissage lui même ne peut pas guider l'exploration.

Nous faisons l'hypothèse que les exemples de contrôle dont nous disposons constituent une exploration raisonnable du système c'est à dire que les actions ont été prises de manière à atteindre l'objectif posé en cherchant à minimiser les risques. Ces exemples ne sont cependant pas supposés avoir été générés par une politique optimale.

8.3 Travaux connexes

Parmi les méthodes consistant à expliciter une fonction de valeur, il semble intéressant de distinguer deux types d'approches. La première est celle des modèles *partiellement observables* (POMDPs) dans lesquels la distinction est faite entre l'espace des observations et celui des états. Dans ces modèles, la décision est prise au regard de la séquence des observations précédentes contrairement au cas *observable* où seule la dernière observation suffit à prendre la décision. Si le formalisme des POMDPs semble plus approprié pour répondre au problème posé, nous nous intéresserons également aux travaux sur les MDPs qui permettent d'aborder des problèmes de grande taille en utilisant une approximation pour représenter la fonction de valeur.

8.3.1 Approximation de la fonction de valeur

L'approximation de la fonction de valeur est un sujet très présent dans la littérature car celle-ci devient nécessaire dès qu'il s'agit d'appliquer la théorie sur l'apprentissage par renforcement à des problèmes réels. Lorsque l'espace d'état est trop grand pour que la fonction de valeur puisse être mémorisée, il est usuel d'utiliser une approximation de la fonction basée sur une représentation plus compacte. Ce problème se pose en particulier si l'espace d'état est continu puisque celui-ci est alors infini. Remarquons à ce sujet que toutes les approches basées sur une discrétisation de l'espace d'états font partie de la grande famille des techniques d'approximation de la fonction de valeur. Parallèlement au problème de la représentabilité de la fonction, notre capacité à explorer les états du système constitue également un facteur limitant qui nous oblige à ne calculer qu'une approximation de la fonction de valeur. L'utilisation d'une représentation compacte de la fonction de valeur nous intéresse alors pour sa capacité à généraliser l'apprentissage.

Il existe de nombreux exemples dans la littérature d'expériences fructueuses d'utilisation d'une approximation de la fonction de valeur pour résoudre des problèmes de grande taille. En 1996 Bertsekas et Tsitsiklis [Bertsekas and Tsitsiklis, 1996] ont proposé un tour d'horizon des méthodes.

Il est possible de représenter la fonction de valeur sous la forme d'un réseau de neurone. Ce type d'approche, appelé *neuro-dynamic programming*, permet d'aborder des problèmes de grande dimension, comme le montre par exemple le travail de Coulom [Coulom, 2002], mais présente l'inconvénient de nécessiter une très grande quantité de données pour converger.

L'écriture de la fonction de valeur comme une combinaison linéaire de fonction support est également très étudiée

$$V_{\vec{w}}(\vec{x}) = \sum_i w_i \phi_i(\vec{x}). \tag{8.18}$$

Remarquons que la discrétisation sur une grille de la fonction de valeur est un cas particulier de cette représentation. Plus généralement les fonctions support peuvent être un codage de l'espace d'état en tuiles de différentes tailles. Il est également usuel d'utiliser des Gaussiennes normalisées. Les fonctions supports peuvent être calculées à partir de caractéristiques propres à l'application visée.

En 2002, Smart a étudié dans sa thèse [Smart, 2002] l'utilisation d'une approximation de la fonction de valeur à base de régression localement pondérée (LWR) pour rendre l'apprentissage par renforcement utilisable en robotique réelle. Un nombre relativement important d'adaptations doivent être apportées à la méthode de régression pour permettre un apprentissage robuste et efficace.

Geist et Pietquin [Geist and Pietquin, 2010] ont récemment présenté une revue des méthodes d'approximation paramétriques de la fonction de valeur à partir d'un ensemble d'échantillons soit *en ligne*, soit *hors ligne*. Cette présentation distingue trois grandes classes de méthodes. La première consiste à traiter le problème de l'apprentissage de la fonction de valeur comme un problème d'apprentissage supervisé en utilisant une estimation des valeurs associées aux points d'apprentissage, calculée à l'aide de l'opérateur de Bellman échantillonné. Les méthodes TD, SARSA et Q-Learning avec approximation de la fonction de valeur sont ainsi présentées comme des descentes de gradient stochastique permettant de réaliser un apprentissage supervisé respectivement de la fonction de valeur pour une politique donnée, de la fonction Q pour une politique donnée et de la fonction Q optimale. La seconde classe d'approches, présentées comme

des méthodes par résidus, consistent à minimiser la distance entre la Q fonction et son image par l'opérateur de Bellman qui, en pratique, est échantillonné. Baird [Baird, 1995] a ainsi proposé des algorithmes de type descente de gradient stochastique utilisant la méthode des résidus. Les méthodes GPTD [Engel *et al.*, 2005] et KTD [Geist *et al.*, 2009] sont fondées sur une minimisation du résidu au sens des moindres carrés. La troisième catégorie de méthodes cherche à minimiser la distance entre la fonction Q et la projection dans l'espace des hypothèses de son image par l'opérateur de Bellman, également échantillonné. LSTD [Bradtke *et al.*, 1996] utilise ainsi la méthode des moindres carrés pour minimiser cette distance. D'autres méthodes comme GTD2 et TDC [Sutton *et al.*, 2009] se fondent sur une descente de gradient stochastique pour atteindre l'objectif. Enfin, en considérant que la composition de l'opérateur de Bellman avec la projection dans l'espace des hypothèse est un opérateur contractant, il existe un unique point fixe qui peut être obtenu en appliquant itérativement ces deux opérateurs. C'est le principe utilisé par l'algorithme Fitted Q, par LSPE [Nedić and Bertsekas, 2002], qui peut être considéré comme une application de Fitted-Q [Ernst *et al.*, 2005] dans le cas d'une paramétrisation linéaire de la fonction, et par (Q-OSP)[Yu and Bertsekas, 2007].

8.3.2 Fitted Q-iteration

L'algorithme *Fitted Q-iteration* nous intéresse particulièrement car il permet d'apprendre *hors ligne* une représentation approchée de la fonction Q. Il consiste à répéter l'opérateur de Bellman sur une représentation approchée de la fonction Q, réestimée à chaque itération à partir de la trajectoire de travail.

Soit $\{s_t, a_t, r_t\}_{1 \leq t \leq T}$ une trajectoire pré-enregistrée générée par une politique stochastique. Le principe général de l'algorithme est de répéter

$$Q_{k+1} = \texttt{Regress}(D_k(Q_k)), \tag{8.19}$$

où `Regress` est une opération de réestimation des paramètres de la fonction représentant Q et où D_k est l'ensemble de couples suivant :

$$D_k(Q_k) = \left\{ \left[(s_t, a_t) \rightarrow r_t + \gamma \max_{a \in \mathcal{A}} Q_k(s_{t+1}, a) \right] \right\}_{1 \leq t < T}. \tag{8.20}$$

Cette algorithme est très général et peut s'adapter à différentes formes d'approximation de la fonction de valeur.

8.4 Utilisation d'un problème jeu

L'utilisation des techniques d'apprentissage par renforcement dans un cadre médical reste peu explorée aujourd'hui. L'approche que nous envisageons ici est donc très exploratoire et nous avons décidé de nous confronter à un problème jeu de manière à pouvoir disposer d'un simulateur nous permettant d'étudier la problématique de modélisation de l'espace d'état avec la possibilité de valider expérimentalement les résultats.

Nous avons utilisé le cas du pendule sur un chariot comme exemple d'apprentissage guidé par un expert. L'idée est d'enregistrer un humain manipulant le chariot pour constituer des trajectoires servant ensuite à l'apprentissage. En nous limitant à l'utilisation de ces trajectoires pour

apprendre la politique, nous nous plaçons dans des conditions réalistes vis à vis de l'exploration du système.

L'objectif de ce problème jeu est d'amener ou de maintenir un pendule en position verticale haute, qui est un état d'équilibre instable. Classiquement le problème consiste à faire pivoter le pendule à l'aide d'un moteur ne disposant pas d'un couple suffisant pour amener directement le pendule en position haute, obligeant le contrôleur à réaliser plusieurs oscillations avant de pouvoir atteindre l'objectif. Une autre variante consiste à partir d'un pendule en position verticale sur un chariot mobile et à déplacer le chariot de manière à maintenir le pendule dans cet état d'équilibre instable. Nous avons utilisé un problème hybride consistant à déplacer un chariot de manière à amener le pendule en position haute et à le maintenir en équilibre. La dynamique du problème utilisée est du premier ordre avec des constantes numériques définies de façon à rendre le simulateur utilisable par un humain.

L'état du système est décrit, comme illustré sur la figure 8.4 par :
– l'abscisse du chariot x,
– la position angulaire du pendule θ,
– la vitesse angulaire du pendule $\dot{\theta}$.

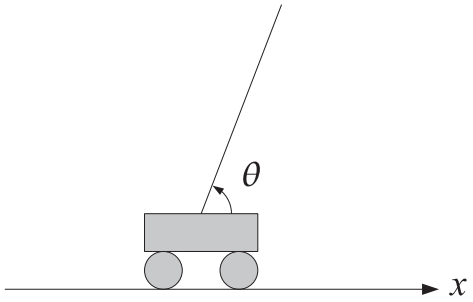


FIGURE 8.4 – Cart-pole balancing

L'objectif du problème est donc d'atteindre et de maintenir la position angulaire du pendule $\theta = \pi/2$. A chaque pas de temps, le chariot peut être déplacé d'un pas à gauche ($a_t = -1$), d'un pas à droite ($a_t = +1$), ou maintenu en place ($a_t = 0$).

La dynamique du système est décrite par les équations suivantes :

$$\begin{aligned} x_{t+1} &= x_t + a_t \\ gravity_t &= sin(\theta_t - \pi/2)/50 \\ push_t &= a_t * sin(\theta_t)/20 \\ friction_t &= \dot{\theta}_t/10 \\ \dot{\theta}_{t+1} &= \dot{\theta}_t + gravity_t + push_t - friction_t \\ \theta_{t+1} &= \theta_t + \dot{\theta}_{t+1} \end{aligned}$$

où $gravity_t$ représente le couple exercé par le champ gravitationnel sur le pendule, $push_t$ l'effort engendré par le déplacement du chariot et $friction_t$ les frottements s'opposant à la rotation du pendule. Les constantes associées aux différents efforts ont été choisies empiriquement de façon à rendre le système contrôlable par un humain sur une simulation cadensée à 20Hz.

La récompense est calculée sur la position angulaire par

$$r_t = sin(\theta_t) \tag{8.21}$$

ce qui nous donne une récompense maximale $max(r_t) = 1$ lorsque $\theta_t = \pi/2$.

La solution du problème étudié nous est donnée par la fonction de valeur optimale que nous pouvons estimer en discrétisant l'espace d'état à l'aide d'une grille fine et en utilisant l'algorithme Q-learning avec une politique $\epsilon - greedy$. Nous avons ainsi construit la fonction de valeur représentée sur la figure 8.5. Nous pouvons noter que certaines zones de la grille (le

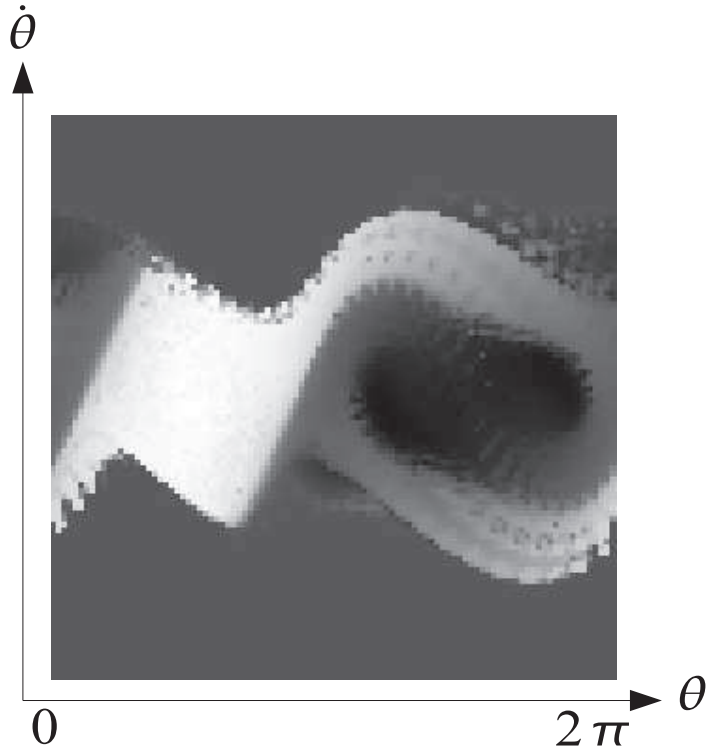


FIGURE 8.5 – Fonction de valeur construite avec Q-learning sur une grille 100×100 avec un facteur d'actualisation $\gamma = 0.95$. Les points clairs correspondent à des états de valeur élevée. La zone grise uniforme correspond à des états non atteignables.

fond gris sombre uniforme) ne sont pas atteignables par le système étant donné ses contraintes dynamiques. Les états correspondants sur la grille n'ont donc aucune utilité.

Une base de trajectoires a été constituée en enregistrant un humain contrôlant le chariot. Cette base est composée de 6500 échantillons du triplet état-action-récompense $(\theta,\dot{\theta},a,r)$. Tous les exemples d'apprentissage présentés dans la suite de ce chapitre sont basés sur cette même base de trajectoire.

8.5 Apprentissage *hors ligne*, à partir de trajectoires

L'algorithme Q-learning [Watkins and Dayan, 1992] permet d'apprendre la valeur optimale $Q^*(s,a)$ des actions suivies sans connaissance sur la dynamique du système. L'apprentissage est réalisé en mettant à jour avec un taux d'apprentissage α la fonction Q après chaque transition

(s_t, a_t, s_{t+1}, r_t) visitée :

$$\begin{aligned} Q(s_t, a_t) &\leftarrow (1-\alpha)Q(s_t, a_t) + \\ &\quad \alpha\left(r_t + \gamma \max_{a'} Q(s_{t+1}, a')\right). \end{aligned}$$

Sous certaines conditions il peut être montré que l'algorithme converge vers Q^*.

Cet algorithme se fonde sur l'opérateur de Bellman (voir paragraphe 8.1.5) pour calculer l'espérance de récompense en séparant d'un côté la récompense immédiate r_t et de l'autre l'espérance de récompense future qui serait obtenue en prenant la meilleure action suivante $\max_{a'} Q(s_{t+1}, a')$. L'espérance est calculée *en ligne* grâce à l'utilisation du taux d'apprentissage α.

Dans le cas d'un apprentissage *hors ligne*, c'est à dire à partir d'une séquence déterminée d'échantillons, il est possible d'utiliser l'algorithme Q-learning en rejouant un grand nombre de fois la même séquence. Cette technique, parfois utilisée, ne nous paraît cependant pas être la plus adaptée puisque l'utilisation d'un taux d'apprentissage n'a pas vraiment de sens si nous travaillons *hors ligne* et pose un certain nombre de problèmes, comme le réglage de sa décroissance. Nous proposons l'algorithme 5 dans lequel l'utilisation du facteur de mise à jour α est remplacée par un calcul explicite de la moyenne sur l'ensemble des visites pour estimer l'espérance de la valeur :

$$Q_{i+1}(s, a) = E\left[R_t + \gamma \max_{a'} Q_i(s_{t+1}, a') \middle| S_t = s, A_t = a\right]. \tag{8.22}$$

Cette opération est répétée jusqu'à la convergence de Q_i. Il est probable que certaines actions n'apparaissent jamais dans certains états ou bien qu'elle ne soient prises qu'un nombre insignifiant de fois. Nous supposons alors que ces actions, qui n'ont pas été explorées par l'expert humain et que nous ne pouvons donc pas évaluer, sont potentiellement dangereuses et ne doivent pas apparaître dans la politique. Les Q-valeurs sont initialisées de façon « pessimiste » et nous utilisons un seuil (th) pour éviter de les mettre à jour si le nombre de visites n'est pas significatif.

L'algorithme que nous proposons s'apparente en fait à l'algorithme *Fitted Q-iteration* exprimé dans le cas d'une approximation tabulaire de la fonction de valeur. La gestion « pessimiste » des actions non explorées est spécifique à notre cadre expérimental.

8.6 Illustration du problème de la représentation de l'espace d'état

Nous étudions ici le problème de l'apprentissage d'une politique sur des grilles de différentes tailles avec la contrainte que l'exploration du système est limitée à la base de trajectoires. Comme mentionné précédemment, travailler avec un ensemble limité de trajectoires implique que l'espace d'état ne sera que partiellement exploré. Ceci peut être illustré en dessinant la fonction de valeur obtenue en utilisant l'algorithme 5 sur une grille 100×100. Etant donné que la grille est composée de 10000 états et que nous ne disposons que de 6500 échantillons, le nombre moyen d'échantillons par état est inférieur à un. Nous pouvons vérifier sur la figure 8.6 que la majorité des états restent inexplorés.

Pour mesurer *a posteriori* la qualité du modèle, nous lançons 40 trajectoires sur le simulateur, de 500 pas de temps chacune, en utilisant la politique issue des Q-valeurs calculées. La moyenne de la récompense reçue est calculée sur les 20000 pas de temps. La position de départ des trajectoires est définie à $\theta = \frac{3\pi}{2}$ (position basse) avec une petite perturbation aléatoire $(-0,5 \leq \dot{\theta} \leq 0,5)$.

Entrées : Une séquence du triplet état-action-récompense (s_t, a_t, r_t) avec $t = 1...T$
$Visits \leftarrow 0$;
pour $t = 1$ *à* $T - 1$ **faire**
  $Visits(s_t, a_t) \leftarrow Visits(s_t, a_t) + 1$;
fin
$Q_0 \leftarrow ValeursInitiales$;
$i \leftarrow 0$;
répéter
  $i \leftarrow i + 1$;
  $Q_i \leftarrow 0$;
  pour $t = 1$ *à* $T - 1$ **faire**
    $q \leftarrow r_t + \gamma max_{a'}(Q_{i-1}(s_{t+1}, a'))$;
    $Q_i(s_t, a_t) \leftarrow Q_i(s_t, a_t) + q$;
  fin
  pour tous les $s, a \in S, A$ **faire**
    si $Visits(s, a) > th$ **alors**
      $Q_i(s, a) \leftarrow \frac{Q_i(s,a)}{Visits(s,a)}$;
    sinon
      $Q_i(s, a) \leftarrow Q_{i-1}(s, a)$;
    fin
  fin
jusqu'à *Convergence de* Q;

Algorithme 5: QD-Iteration : Itération sur les valeurs de Q, calculées à partir d'un ensemble fini de transitions.

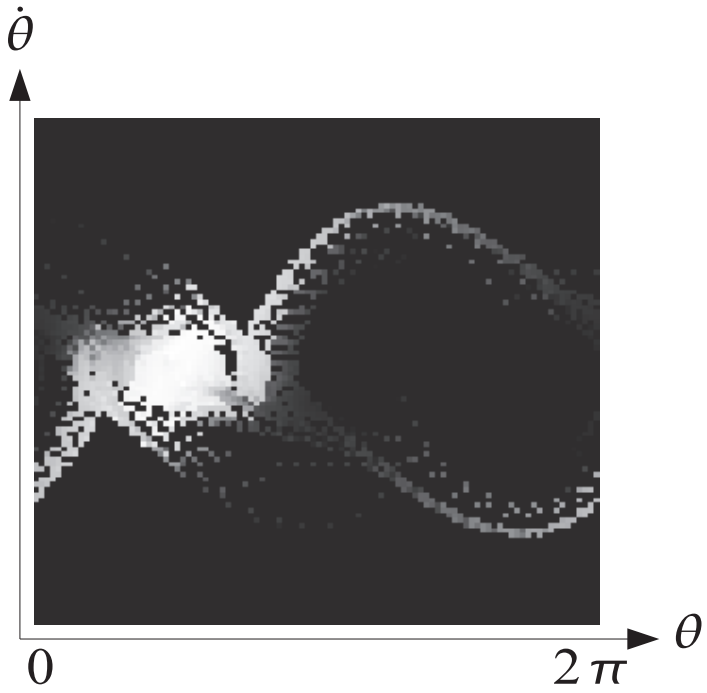


FIGURE 8.6 – Fonction de valeur calculée sur une grille 100×100 en utilisant QD-Iteration avec un facteur d'affaiblissement $\gamma = 0.95$. Le fond, gris sombre, correspond à des états non visités.

Avec 10000 états, la récompense moyenne reçue est inférieure à 0 (≈ -0.3) ce qui signifie que le contrôleur n'arrive pas à amener le pendule en position verticale. Le découpage est manifestement trop fin et le modèle n'a pas de capacité de généralisation. Comment l'algorithme d'apprentissage se comporte-t-il sur des grilles de différentes tailles ? La figure 8.7 montre la fonction de valeur pour une grille composée de 25 états. Sur cette grille 5×5, la fonction de valeur semble couvrir une zone plus importante de l'espace d'état (en comparaison avec la grille 100×100). Cependant la grille à 25 états ne permet pas de produire un contrôleur efficace. La récompense moyenne reçue avec ce modèle est également négative (≈ -0.6).

Nous avons utilisé l'algorithme 5 sur différentes tailles de grilles. La récompense reçue par le contrôleur résultant de l'apprentissage est tracée sur la figure 8.8. Les meilleures politiques ont été obtenues pour les grilles de tailles 64 (0.62), 169 (0.85) et 289 (0.76). Ces résultats illustrent le fait que le nombre d'états doit être choisi en fonction du nombre d'échantillons d'apprentissage. Nous faisons face à un dilemme précision - capacité de généralisation. Des modèles simples ont de meilleures capacités de généralisation avec le coût d'une précision plus faible. Il est intéressant de noter la forme générale de la courbe sur la figure 8.8, composée d'une croissance brutale pour atteindre rapidement une zone maximum suivie par une décroissance lente au fur et à mesure que le modèle se complexifie. Le principe d'Ockham semble s'appliquer ici. Les meilleurs modèles sont les plus simples permettant d'expliquer nos observations. Relativement peu de modèles en forme de grilles permettent d'atteindre l'objectif et nous ne découvrons la qualité du modèle qu'après exécution. Notre objectif ici étant de choisir un espace d'état pour faire de l'apprentissage sur des trajectoires, nous avons besoin d'un critère permettant d'évaluer le modèle *a priori*, c'est à dire en ne se basant que sur les trajectoires.

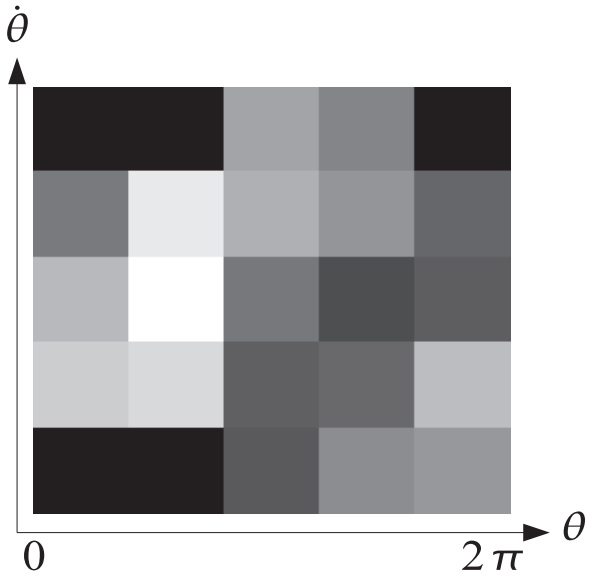


FIGURE 8.7 – Fonction de valeur calculée sur une grille 5×5 en utilisant QD-Iteration avec un facteur d'actualisation $\gamma = 0.95$.

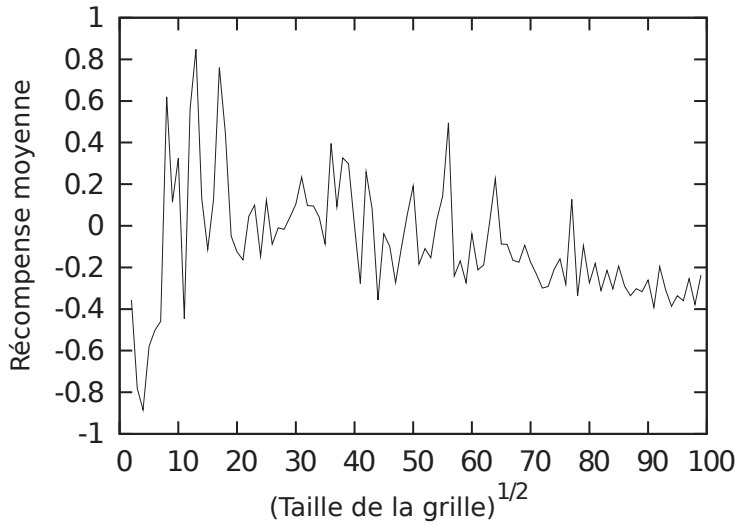


FIGURE 8.8 – Récompense moyenne reçue par le contrôleur construit avec QD-Iteration sur différentes tailles de grilles avec un facteur d'actualisation $\gamma = 0.95$.

8.7 Approche proposée

Nous proposons d'utiliser la vraisemblance du modèle comme critère permettant de construire une représentation de l'espace d'états efficace pour l'apprentissage du contrôle.

8.7.1 Le critère de vraisemblance

La vraisemblance des paramètres d'un modèle λ pour une séquence d'observation O donnée est définie par :

$$\mathcal{L}(\lambda|O) = P(O|\lambda) \tag{8.23}$$

Le critère de maximum de vraisemblance est habituellement utilisé pour estimer les paramètres de modèles stochastiques tels que les modèles de Markov cachés ou, plus généralement, pour estimer les paramètres d'un réseau bayésien dynamique (DBN).

Chrisman a proposé d'utiliser un modèle prédictif de type processus décisionnel de Markov partiellement observable (POMDP) pour gérer le recouvrement de perceptions sur des problèmes discrets d'apprentissage par renforcement [Chrisman, 1992]. Dans cette approche le POMDP, estimé par maximum de vraisemblance, permet d'avoir un espace d'état plus grand que l'espace des observations. McCallum a repris cette approche en proposant un critère d'utilité pour choisir où diviser les états du modèle [McCallum, 1992]. Sa méthode consiste à comparer l'espérance de récompense dans un état en fonction des différentes transitions entrantes et à diviser les états dans lesquels la récompense varie significativement en fonction de l'état précédent.

Nous nous intéressons ici à la possibilité d'utiliser le critère de maximum de vraisemblance pour apprendre l'espace d'état sous la forme d'un POMDP. L'intuition à l'origine de cette proposition est probablement la même que pour l'approche proposée par Chrisman qui est qu'un modèle qui « explique bien » le processus observé fournira une bonne représentation pour apprendre une politique de contrôle. Lors du processus de maximisation de la vraisemblance pour un HMM, le modèle tend en effet à devenir « plus déterministe ». Le fait de minimiser l'incertitude sur les transitions nous permet de mieux prévoir l'effet des actions sur le système. Il en va de même pour le modèle de récompenses : dans le cas de densités gaussiennes, des vraisemblances plus élevées pourront être atteintes par des gaussiennes à plus petite variance et nous pouvons remarquer que lors du processus de maximisation, les variances des gaussiennes tendent effectivement à diminuer. Un modèle « plus déterministe » (compatible avec les observations) aura ainsi une vraisemblance plus élevée.

8.7.2 Apprentissage de l'espace d'état avec un DBN

Les réseaux bayésiens dynamiques sont un formalisme général pour représenter des indépendances conditionnelles dans des processus stochastiques. Les algorithmes d'inférence et d'apprentissage associés sont applicables à différentes structures de HMMs qui ne sont que des cas particuliers de DBNs, comme cela a été montré par [Murphy, 2002]. Notre modèle a la forme d'un POMDP et est défini par :

- une variable d'état S (discrète),
- une variable observée O (continue),
- une variable de contrôle A (discrète),
- une variable récompense R (continue),

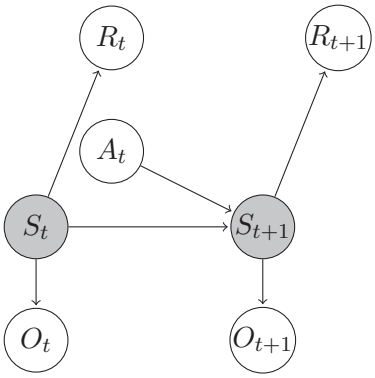


FIGURE 8.9 – Représentation graphique du POMDP. Les variables de contrôle et d'états A et S sont discrètes alors que l'observation O et la récompense R sont des nœuds continus gaussiens.

- une distribution de probabilité initiale $\Pi = P(S_0)$
- une fonction de transition stochastique $T = P(S_t|S_{t-1}, A_{t-1})$,
- une fonction d'observation stochastique $O = P(O_t|S_t)$,
- une fonction de récompense stochastique $R = P(R_t|S_t)$.

La figure 8.9 donne la représentation graphique du modèle. Dans notre exemple du pendule sur le chariot, la fonction d'observation est une gaussienne à deux dimensions définie sur $O = (\alpha, \dot{\alpha})$ et la fonction de récompense est une gausienne de dimension un.

Notre base d'apprentissage contient des séquences du triplet observation-action-récompense (o_t, a_t, r_t). Etant donné un nombre d'états N ($S \in (1, \ldots, N)$) fixé arbitrairement, nous pouvons apprendre les paramètres du modèle (fonction de transition, fonction d'observation et fonction de récompense) en maximisant sa vraisemblance avec un algorithme EM.

L'algorithme *Expectation-Maximization* consiste en une alternance de deux sous-procédures, la première pour calculer les beliefs $b_{1..T}$ ($b_t = P(X_t|a_{1..T}, o_{1..T}, r_{1..T}, \lambda)$) et la seconde pour estimer les paramètres du modèle $\lambda' = (\Pi', T', O', R')$ d'après $b_{1..T}$, $a_{1..T}$,$o_{1..T}$ et $r_{1..T}$. L'algorithme produit une séquence de modèles $\lambda_0 \rightarrow \lambda_1 ... \rightarrow \lambda_K$ jusqu'à la convergence de la vraisemblance $P(a_{1..T}, o_{1..T}, r_{1..T}|\lambda)$. L'algorithme nécessite la définition d'un premier modèle λ_0 pour initier la procédure. Nous avons donc utilisé l'algorithme K-means pour effectuer une première partition de l'ensemble des observations O en N classes à partir desquels nous pouvons estimer des distributions initiales pour O et R. Dans le modèle initial T et Π sont choisis uniformes. Remarquons cependant que l'algorithme des K-means est aussi un algorithme de type E.M. nécessitant une initialisation et nous choisissons N échantillons aléatoirement comme premiers centres. Ce choix aléatoire nous conduit au final à différents points de convergences possibles pour le modèle puisque E.M. ne donne qu'un maximum local.

8.7.3 Construction de la politique

Suite à l'apprentissage des paramètres du DBN à partir des trajectoires, nous disposons d'un modèle de transitions $T = P(S_t|S_{t-1}, A_{t-1})$ et d'un modèle de la fonction de récompense $R = P(R_t|S_t)$. La fonction de récompense est définie par un ensemble de distributions gaussiennes. L'espérance de la récompense dans un état est donc donnée par la moyenne de la distribution gaussienne correspondante. Ces éléments nous permettent donc de calculer avec l'algorithme Value-Iteration la fonction de valeur optimale (pour ce modèle). Nous avons donc expérimenté

cette première façon de construire la politique qui a donné de bons résultats pour certains modèles. Cependant, nous avons globalement obtenu de meilleurs résultats en construisant les Q-valeurs à partir des vrais récompenses contenues dans les trajectoires d'apprentissage. Ceci est probablement dû à la perte d'information lors de la modélisation de la récompense.

Pour l'apprentissage de la politique dans le POMDP estimé, Chrisman a utilisé une adaptation de Q-learning en utilisant l'approximation suivante : $Q(\mathbf{b}, a) \approx \mathbf{Q_a}.\mathbf{b}$ où $\mathbf{b}$ est un belief-state et $\mathbf{Q_a}$ est le vecteur des valeurs pour l'action a. L'utilisation d'un vecteur unique par action ne permet pas de représenter la fonction de valeur optimale. Cependant cette approximation est simple et s'avère efficace. Elle a été étudiée sous le nom de « Replicated Q-learning » dans [Littman *et al.*, 1995]. Cette approximation, qui revient à se fonder sur le MDP sous-jacent pour estimer les valeurs de Q, est aussi connue sous le nom d'approximation Q-MDP [Cassandra, 1998]. Après chaque action a, les valeurs du vecteur $\mathbf{Q_a}$ sont mises à jour proportionnellement au belief b en utilisant :

$$\Delta Q_a(s) = \alpha b(s)(r + \gamma max_{a'} Q_{a'}(b') - Q_a(s))$$

Pour adapter QD-Iteration à l'utilisation d'un belief-state nous avons testé cette façon de mettre à jours les Q-valeurs proportionnellement au belief. Il s'avère cependant que les résultats étaient légèrement meilleurs en ne mettant à jour que l'état le plus probable à chaque pas de temps (donné par $s_t = argmax_s b(s)$). Nous avons gardé l'approximation $Q(\mathbf{b}, a) \approx \mathbf{Q_a}.\mathbf{b}$ pour calculer l'espérance de la récompense, ce qui conduit à l'algorithme 6. Les beliefs associés aux trajectoires de la base sont calculés par inférence dans le DBN.

Entrées : Une séquence du triplet belief-action-récompense (b_t, a_t, r_t) avec $t = 1...T$
$Visits \leftarrow 0$;
pour $t = 1$ *à* $T - 1$ **faire**
  $s_t \leftarrow argmax_s(b_t(s))$;
  $Visits(s_t, a_t) \leftarrow Visits(s_t, a_t) + 1$;
fin
$Q_0 \leftarrow InitialValue$;
$i \leftarrow 0$;
répéter
  $i \leftarrow i + 1$;
  $Q_i \leftarrow 0$;
  pour $t = 1$ *à* $T - 1$ **faire**
    $MaxQ \leftarrow max_{a'}(\sum_s Q_{i-1}(s, a')b_{t+1}(s))$;
    $Q_i(s_t, a_t) \leftarrow Q_i(s_t, a_t) + (r_t + \gamma MaxQ)$;
  fin
  pour tous les $s, a \in S, A$ **faire**
    si $Visits(s, a) > th$ **alors**
      $Q_i(s, a) \leftarrow \frac{Q_i(s,a)}{Visits(s,a)}$;
    sinon
      $Q_i(s, a) \leftarrow Q_{i-1}(s, a)$;
    fin
  fin
jusqu'à *Convergence de* Q;

Algorithme 6: QD-Iteration avec des belief-states

Nous utilisons également l'approximation $Q(\mathbf{b}, a) \approx \mathbf{Q_a}.\mathbf{b}$ pour construire le contrôleur. La politique associée est donnée par $\pi(\mathbf{b}) = argmax_a \mathbf{Q_a}.\mathbf{b}$.

8.7.4 Résultats expérimentaux

La figure 8.10 montre la fonction de valeur obtenue en utilisant QD-Iteration sur un DBN à 25 états. Ce résultat est à rapprocher de la fonction de valeur calculée sur la grille 5×5 représentée sur la figure 8.7. Nous pouvons remarquer que les états ne sont pas uniformément répartis dans l'espace. La zone proche de la position but ($\Pi/2$) a une distribution d'états plus dense que le reste de l'espace.

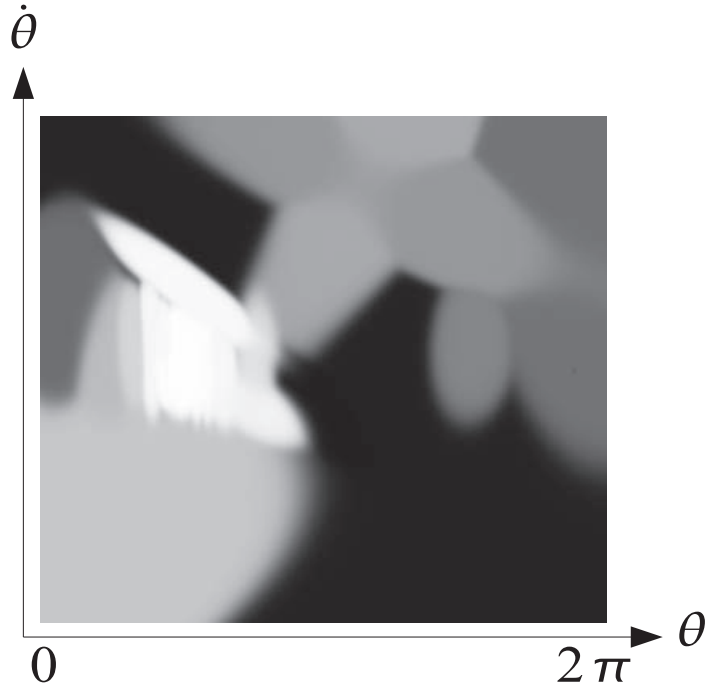


FIGURE 8.10 – Fonction de valeur pour un DBN à 25 états calculée en utilisant QD-Iteration avec des beliefs et un facteur d'actualisation $\gamma = 0.95$.

Comme nous l'avons déjà évoqué, E.M. converge vers un minimum local qui dépend du modèle initial λ_0 utilisé (le modèle de départ dépend lui même d'un tirage aléatoire d'échantillons pour les K-means). Nous avons donc évalué la qualité de la politique obtenue en relançant plusieurs fois l'apprentissage des paramètres à partir de différentes situations initiales. Pour chaque modèle, nous apprenons la politique, toujours en utilisant l'algorithme 6, que nous évaluons expérimentalement en calculant la récompense moyenne $\bar{r}$ obtenue par le contrôleur sur 20000 pas de temps. La variabilité des résultats obtenus s'avère importante comme l'illustre la figure 8.11. Une récompense moyenne $\bar{r} < 0$ signifie que le contrôleur ne parvient pas à amener le pendule en position verticale. Lorsque la récompense moyenne est positive mais proche de 0, cela signifie que le pendule passe la plus grande partie du temps dans la zone supérieure ($\alpha \in [0, \Pi]$) mais que le contrôleur ne parvient pas à le stabiliser. Enfin la récompense moyenne est proche de 1 lorsque le contrôleur parvient à stabiliser le pendule en position verticale. La figure 8.12 montre l'évolution

des résultats obtenus pour différentes tailles de modèles. Le tracé suit approximativement celui de la figure 8.8, caractérisé par une croissance initiale rapide suivie par une décroissance lente.

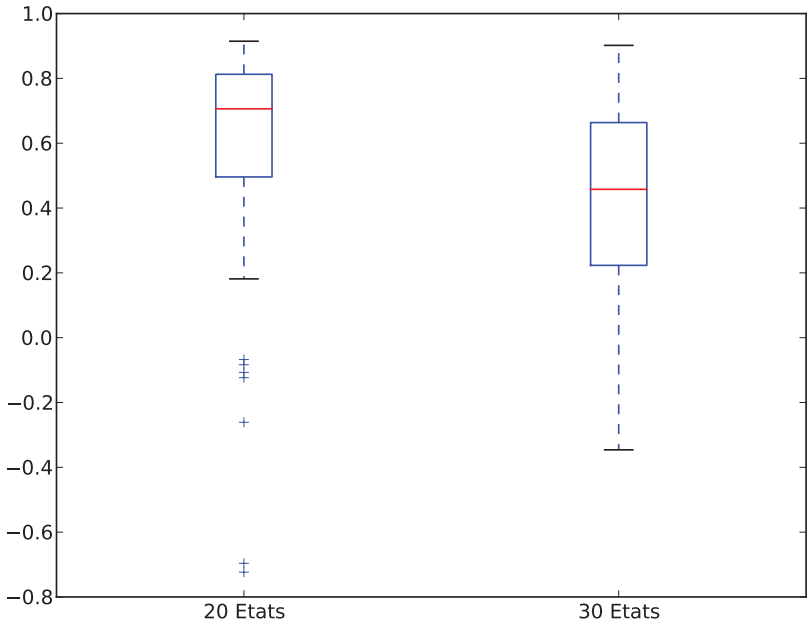


FIGURE 8.11 – Dispersion des récompenses moyennes obtenues pour 60 exécutions de l'algorithme d'apprentissage pour des modèles à 20 et 30 états.

Il n'est pas surprenant que différentes exécutions de E.M. aboutissent à des comportements différents pour le contrôleur puisque l'algorithme converge vers des minima locaux différents. Cependant, en rappelant que notre problème de départ est d'apprendre un contrôle uniquement à partir d'un ensemble fini de trajectoires, une question importante est de pouvoir mesurer la qualité du modèle sans avoir effectivement besoin de l'exécuter. Nous utilisons le critère de vraisemblance pour apprendre les paramètres du modèle. Ce critère est-il prédictif de la qualité du contrôleur ?

Pour répondre à cette question, nous avons comparé la vraisemblance à la récompense moyenne obtenue par le contrôleur pour différents modèles. La vraisemblance du modèle a bien un impact sur la récompense reçue par le contrôleur. La figure 8.13 montre la dispersion de la récompense moyenne obtenue pour deux groupes de modèles à 20 états et deux groupes de modèles à 30 états. Les modèles ont été répartis dans les groupes en fonction de leur vraisemblance. Le premier groupe est composé des 30 modèles ayant les vraisemblances les plus basses alors que le second est composé des 30 modèles ayant les vraisemblances les plus élevées. Nous avons pu montrer par un test de Mann-Whitney que la récompense moyenne était significativement différente entre les deux groupes ($p = 0.029$ pour les modèles à 20 états et $p = 0.008$ pour les modèles à 30 états).

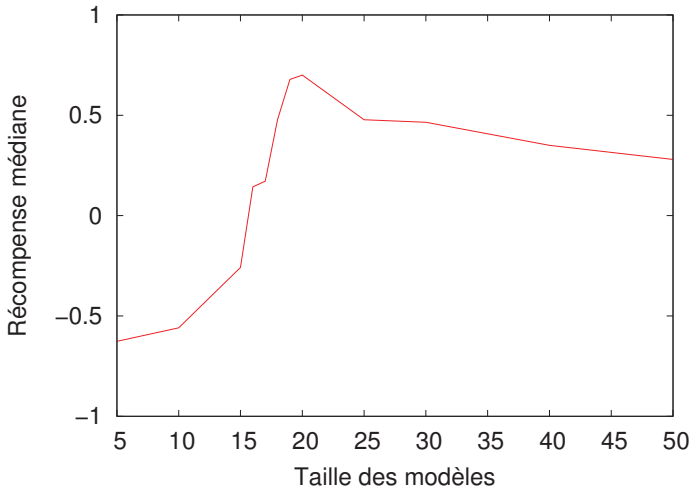


FIGURE 8.12 – Evolution de la récompense obtenue par le contrôleur pour différentes tailles de modèles. La récompense tracée est la médiane des récompenses, calculées pour chaque taille de modèle sur 15 exécutions de l'algorithme d'apprentissage.

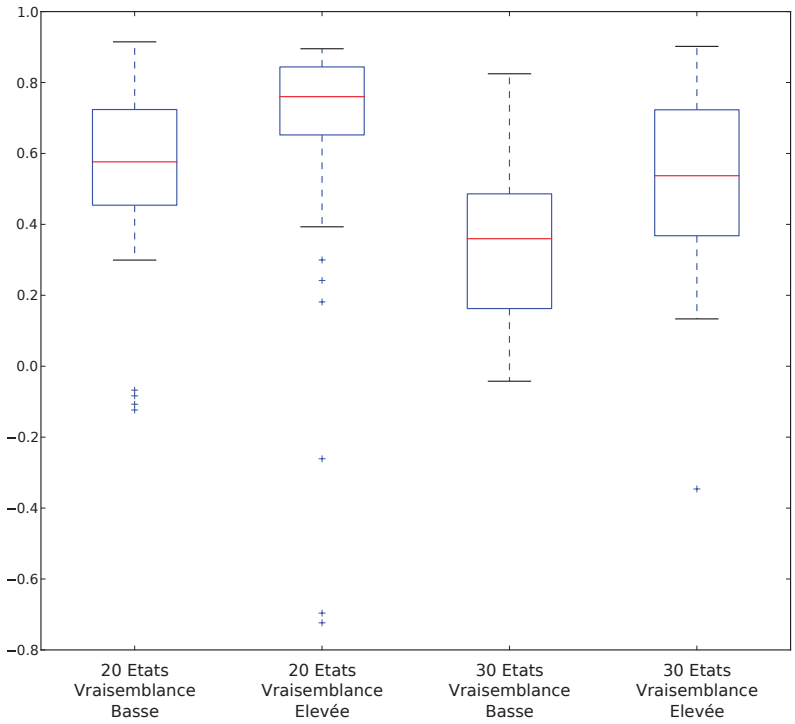


FIGURE 8.13 – Dispersion de la récompense moyenne pour 4 groupes de 30 modèles constitués en fonction de leur vraisemblance.

Discussion

La méthode présentée donne des résultats variables d'un modèle à l'autre en fonction de l'initialisation de l'algorithme d'apprentissage des paramètres (E.M.) mais également en fonction du nombre d'états. Nous avons pu confirmer expérimentalement que la vraisemblance du modèle avait un impact sur la qualité de la politique apprise. Ce critère peut donc être utilisé en pratique pour choisir après plusieurs exécutions de E.M. le modèle qui aura le plus de chance de donner une bonne politique. Notons que la vraisemblance ne peut être utilisée que pour comparer entre eux des modèles ayant le même jeu de paramètres et donc le même nombre d'états. L'évolution des résultats en fonction du nombre d'états confirme la nécessité d'adapter la complexité du modèle au problème. Le critère BIC (Bayesian Information Criterion) [Schwarz, 1978], qui peut être vue comme une expression baysienne du principe de parcimonie, permet de comparer entre eux des modèles de dimensions variables.

8.8 Conclusion

L'approche que nous proposons consiste à apprendre de manière non supervisée une représentation efficace des états du système en maximisant la vraisemblance du modèle. Le modèle dans son expression la plus simple a une structure de POMDP avec une fonction d'observation continue. Cependant, le formalisme des DBN nous permet une souplesse de représentation, ce qui rend l'approche adaptable à diverses situations. Nous pouvons de la même façon traiter le cas d'observations hybrides.

Cette représentation est ensuite utilisée pour apprendre une politique de contrôle maximisant une fonction de récompense qui est une représentation objective d'un but à atteindre. L'algorithme d'apprentissage proposé repose sur l'approximation Q-MDP et est adapté au cas de l'apprentissage *hors ligne* à partir de trajectoires.

Nous avons évalué notre approche sur une variante du pendule sur le chariot. Le problème de discrétisation traité dans le cadre de ce problème jeu est une illustration du problème de la recherche d'un espace d'état caché que nous avons traité sous la contrainte d'une exploration limitée à une trajectoire préenregistrée. Nous avons pu, sur cet exemple, mettre en évidence l'existence d'une relation entre la vraisemblance du modèle et la qualité de la politique apprise.

L'approche proposée permet d'apprendre automatiquement une représentation sur un espace d'état discret qui « explique », au sens de la vraisemblance mathématique, la dynamique des phénomènes observés et les actions prises par le médecin. Le problème principal avec l'apprentissage non supervisé, comme nous l'avons souligné dans le chapitre 5, est qu'il n'est pas possible d'imposer une sémantique à une variable complètement cachée. Nous ne sommes donc pas capable d'interpréter la signification des différents états de notre modèle. Cependant, le fait de poser le problème comme un problème de contrôle nous permet d'associer à chaque état une action thérapeutique, dont le sens est bien connu. Ceci nous permet d'utiliser cette approche pour fournir une recommandation ou générer des alertes, ce qui est le but recherché. L'espace d'état que nous apprenons est, en quelques sortes, une représentation symbolique cachée, entre des perceptions et des décisions.

La principale limite de notre approche est liée à la question de l'évaluation du modèle et de la politique construite. L'algorithme E.M. converge vers un maximum local et la qualité de la politique construite est très variable selon le point de départ de l'apprentissage. Nous avons

mis en évidence que la vraisemblance du modèle était un indicateur de la qualité de la politique mais elle est loin d'en être une mesure. Nous débouchons donc à l'issue de ce chapitre sur une nouvelle problématique, celle de l'évaluation *a priori*, c'est à dire *hors ligne*, de la qualité de la politique apprise. Différentes pistes pourront être explorées pour répondre à cette question. Une première intuition que nous avons est de s'intéresser à la fois à la vraisemblance du modèle et à la valeur de la politique apprise dans le but de maximiser l'espérance de la valeur de la politique sachant les observations. Certaines réponses pourront peut être être trouvées du côté de la théorie de l'information. Comme nous l'avons mentionné dans ce chapitre, un modèle plus déterministe permet de mieux prédire le résultat des actions et donc d'aboutir à des politiques de plus grande valeur. L'entropie de Shannon, qui permet de mesurer une quantité d'information, est probablement un critère intéressant à étudier pour comparer les modèles appris avec l'idée qu'un modèle avec une entropie plus petite serait plus prédictif.

Conclusion

Nous avons abordé dans cette troisième et dernière partie du mémoire le problème de l'apprentissage du modèle à partir d'exemples. Nous avons tout d'abord présenté le principe de l'identification du modèle par maximisation de sa vraisemblance. La vraisemblance du modèle permet d'évaluer son adéquation avec les données d'apprentissage. L'algorithme E.M., présenté dans le cas général de l'apprentissage des paramètres d'un DBN, permet de rechercher, pour une structure donnée, un modèle qui « explique bien » les observations. La méthode E.M. permet ainsi de traiter des problèmes de classification non supervisés en regroupant les données homogènes. Il n'est cependant pas possible de préserver la sémantique des valeurs d'une variable lorsque celle-ci n'est jamais observée. C'est pourquoi l'identification de modèles ne peut pas être traitée de manière complètement non supervisée.

Les chapitres 6 et 7 présentent deux exemples d'apprentissage de modèles à partir de bases de données étiquetées. Lorsque suffisamment de données sont disponibles pour l'apprentissage, la principale difficulté reste la grande variabilité des observations liée à la diversité interindividus et à l'observabilité partielle des phénomènes. En prenant en compte à la fois la dynamique du processus et l'incertitude sur les observation, l'approche bayesienne nous permet cependant d'aboutir à des diagnostics relativement robustes. Nous avons montré dans le chapitre 7 qu'il était possible de spécialiser le modèle pour des classes de patients et d'apprendre cette spécialisation de manière non supervisée pour améliorer la précision des résultats. Une étape de pré-traitement des données, reposant sur une expertise humaine, a permis de faire ressortir l'information pertinente.

Nous avons enfin étudié, dans le chapitre 8, la possibilité de faire de l'apprentissage par renforcement pour élaborer un système de télésurveillance médicale. Il est en effet souvent possible de considérer l'action du médecin comme un problème de contrôle dans lequel les actions thérapeutiques ont pour but d'amener le patient dans un état de bonne santé. La stratégie de traitement peut alors prendre la forme d'une politique dans un processus de décision markovien. Si cette approche nous paraît particulièrement prometteuse en raison de la possibilité d'apprendre la politique à partir de l'observation du médecin dans sa pratique normale, les spécificités du cadre applicatif posent un certains nombre de contraintes auxquelles nous nous sommes intéressés. Nous avons en particulier montré l'importance de choisir une bonne représentation du problème, et proposé d'apprendre de manière non supervisée cette représentation sous la forme d'un DBN par maximisation de la vraisemblance. La principale difficulté de l'approche est liée à la question de l'évaluation du système. En partant du principe qu'il n'est pas possible de « tester » le modèle appris, nous avons besoin de définir un critère d'évaluation du système au regard des données d'apprentissage. Nous nous sommes intéressé à la vraisemblance du modèle et nous avons pu montrer sur un problème jeu que les modèles présentant une plus grande vraisemblance permettaient, en moyenne, d'apprendre une meilleure politique. Ce critère est donc un premier élément de réponse même si celui-ci ne porte pas sur la pertinence politique elle-même.

Conclusion

L'augmentation récente de la production de données numériques, qu'elles soient issues d'appareils médicaux en milieu hospitalier fournissant une grande quantité d'informations en temps réel sur le déroulement d'un traitement ou bien qu'elles proviennent de relevés télétransmis depuis le domicile du patient, nous pousse à faire évoluer nos capacités de traitement. Cela pose des questions d'ordre technologique et scientifique. C'est dans ce contexte que s'inscrit ce travail de thèse. Malgré une évolution scientifique, la médecine moderne reste un art, par essence difficile à modéliser. La richesse de ce domaine ainsi que la dominance de l'esprit humain, qui demeure la référence unique, en font un sujet test pour l'intelligence artificielle.

Le problème de l'analyse de données médicales est, en effet, difficile en raison de leur grande variabilité liée à la complexité des phénomènes du vivant et à leur observabilité partielle. Nous avons montré dans ce travail de thèse, avec différentes expériences applicatives, que la modélisation stochastique permet d'aborder le problème de l'analyse de données médicales avec un formalisme unique. Différentes approches ont été proposées pour construire le modèle.

La première approche envisagée est celle de la modélisation explicite de la démarche médicale qui consiste à raisonner directement sur les règles de diagnostic. Comme nous l'avons vu dans le chapitre 3, cette approche permet de traiter des problèmes dans lesquels les sorties attendues pour le système sont difficiles d'accès soit en raison d'un manque de données soit en raison de l'impossibilité de mesurer ces sorties. Nous avons montré dans le chapitre 4 que le formalisme des DBN permet de traiter efficacement des problèmes d'inférence de grande taille en s'appuyant sur la structure du problème. La modélisation de la démarche présente cependant l'inconvénient de nécessiter beaucoup de temps d'expert et d'aboutir à un résultat peu évolutif.

Le second axe de travail de cette thèse est celui de l'apprentissage automatique du modèle. Cet apprentissage peut tout d'abord se faire de manière supervisée lorsque suffisamment de données étiquetées sont disponibles. Nous avons présenté dans les chapitres 6 et 7 deux exemples d'applications réelles dans lesquelles l'apprentissage automatique des paramètres du modèle a permis d'aboutir à des résultats robustes malgré la grande variabilité des mesures. L'apprentissage automatique permet notamment d'envisager la construction de modèles plus spécifiques avec la constitution de sous-groupes de patients comme illustré dans le chapitre 7. Cependant, l'expérience montre qu'il est important de pouvoir travailler sur l'information la plus pertinente possible. Une étape de traitement préalable des données, reposant sur une expertise humaine, est nécessaire pour aboutir à une bonne reconnaissance des situations. Une autre approche de l'apprentissage que nous avons envisagé est celle de l'apprentissage par renforcement à base d'exemples. Cette approche rend possible, par l'observation du résultat des actions prises par le passé, l'apprentissage d'une politique dans le but de normaliser à long terme un ou plusieurs paramètres. Cette démarche se heurte au problème du choix d'une représentation de l'espace d'état adaptée à la quantité de données disponibles pour apprendre efficacement une politique. Nous avons montré dans le chapitre 8 qu'il était possible d'apprendre automatiquement une telle représentation en utilisant le critère de la vraisemblance. Une mise à l'épreuve de cette technique dans le cadre d'une application réelle nous permettra d'évaluer son intérêt dans la pratique.

Une difficulté inhérente à la thématique médicale est celle de la validation des modèles. La principale question à laquelle nous n'arrivons pas encore bien à répondre est celle de la pertinence *a priori* du modèle construit. Est-il possible de mesurer l'efficacité du modèle dans un contexte où peu de données sont disponibles et où la mesure de l'état exact du patient sur ces données n'est pas possible ? Le problème est d'abord théorique car nous avons besoin de mettre en place

une évaluation objective dans un contexte incertain. Nous nous sommes intéressés dans ce travail à la notion de vraisemblance du modèle. C'est en effet un critère qui renseigne sur les éventuelles contradictions entre le modèle et les données dont nous disposons. D'autres éléments de réponse pourront éventuellement être trouvés dans le champ de la théorie de l'information puisque nous souhaitons évaluer et minimiser l'incertitude liée au modèle construit.

L'étape de traitement préalable et de sélection des données reste, dans notre expérience, réalisée de manière *ad hoc* en tirant parti de l'expertise du spécialiste. Ce travail, réalisé en dehors de la modélisation stochastique, est indispensable pour extraire les caractéristiques pertinentes des signaux en entrée. La recherche de l'automatisation de cette étape de pré-traitement autoriserait probablement une avancée dans l'interprétation automatique de données médicales. Ce travail nécessite de pouvoir évaluer objectivement la fonction apprise, par exemple à partir d'une base de données étiquetées. Une piste de travail pourrait être la recherche d'une combinaison d'opérateurs élémentaires à partir de méthodes évolutionnaires.

Dans notre expérience, les systèmes construits autour des modèles que nous avons présentés ont reçu une bonne appréciation de la part des patients. En ce qui concerne les médecins, il convient de distinguer l'avis favorable de ceux qui ont participé à la construction de ces systèmes de celui des praticiens qui n'ont pas pris part à cette élaboration et ont plutôt vécu l'arrivée de telles solutions comme une charge et une responsabilité supplémentaires. Il existe en effet un certain nombre de verrous au développement de la télésurveillance médicale. Un premier obstacle, est la nécessité pour les patients bénéficiaires de ces systèmes de disposer d'un accès internet et du matériel informatique nécessaire au recueil des données. La mise en place d'une connexion, le prêt d'un ordinateur ou d'un autre terminal adapté au patient, l'organisation logistique nécessaire au bon fonctionnement de la télésurveillance médicale ainsi que le développement et l'hébergement du service ont un coût qui, même s'il est faible devant le coût de la prise en charge des pathologies chroniques, doit être pris en compte par les organismes de soins. Les évolutions récentes de la réglementation des systèmes de remboursement vont dans le sens de la reconnaissance et de la prise en charge de l'acte de télémédecine. Des difficultés organisationnelles peuvent également apparaître pour les professionnels de santé utilisateurs de ces systèmes. Celles-ci sont liées à l'arrivée d'un outil nouveau dans la pratique quotidienne. Même avec les méthodes de filtrage que nous développons, ces systèmes engendrent une augmentation de la quantité d'information à traiter par les médecins. Si le recours à l'outil informatique permet d'accéder plus efficacement aux données du patient, par exemple au moment de la consultation, l'utilisation d'un écran, comme élément de la *pratique* médicale, est un changement important qu'il faut apprendre à gérer. La question de l'ergonomie des solutions proposées et de leur adéquation avec la pratique médicale doit être traité. L'évaluation *a posteriori* de leur efficacité permettra de les améliorer. C'est en faisant la preuve de leur efficacité que ces systèmes pourront trouver l'adhésion des professionnels de santé, la question de la responsabilité médicale restant centrale. Disposer d'un suivi en continu des patients induit nécessairement une responsabilité nouvelle qui ne peut être acceptée par le médecin si celui-ci ne dispose pas d'une vision claire de la pertinence du système. Enfin, l'informatisation et le partage de données médicales soulèvent des questions d'ordre déontologique auxquelles les systèmes de télésurveillance doivent répondre pour recevoir l'adhésion des patients et des professionnels de santé. Le cadre d'utilisation des données doit être clairement défini et le respect de la confidentialité doit être garanti. Un obstacle est que les tiers technologiques, contrairement aux professionnels de santé, ne sont pas toujours suffisamment sensibilisés aux questions liées aux droits des patients usagers. La reconnaissance de la notion d'hébergeur de données de santé dans la loi de 2002 et la mise en place, dans le décret du 4 janvier 2006, d'une

procédure d'agrément de ces hébergeurs est un élément de réponse qui permettra de donner des garanties techniques et organisationnelles aux usagers de ces systèmes. Le cadre légal pose des limites à l'utilisation des données permettant de garantir les droits des patients. Ces limites, nécessaires, pourront cependant être parfois vécues comme un frein au développement de ces systèmes fondés sur la circulation et le partage de l'information.

Malgré l'existence de ces verrous, la télésurveillance médicale ne peut que se développer. L'enjeu est important et de nombreux programmes régionaux et nationaux visent à soutenir et à organiser le développement de la télésanté. Les méthodes et les expériences proposées dans cette thèse sont toutes fondées sur un même formalisme. Il ressort de ce travail qu'une partie importante de la méthodologie utilisée pour réaliser une application de télésurveillance médicale peut être indépendante de la pathologie considérée. L'apprentissage automatique de la fonction de filtrage et l'utilisation de méthodes d'évaluation objectives permettront de développer une méthodologie générique et rationnelle de conception de systèmes d'aide à la décision. La télésurveillance des maladies chroniques n'est pas le seul domaine pouvant bénéficier de tels développements. L'aide au maintien à domicile de personnes en perte d'autonomie est un domaine ou les enjeux sont importants. La population générale est également concernée puisqu'elle pourra bénéficier de systèmes visant à aider les personnes à rester en bonne santé, à assurer un bon usage des médicaments ou encore à prévenir le développement de maladies chroniques par la mise en place de mécanismes automatiques de détection précoce.

Travailler à l'interface de plusieurs disciplines est particulièrement enrichissant et les deux disciplines peuvent en bénéficier scientifiquement. D'un côté, le développement d'applications de télésurveillance médicale soulève de nouvelles problématiques théoriques qui viennent alimenter la recherche et faire progresser la connaissance en informatique. De l'autre, la découverte de nouvelles techniques d'analyse des données permettra probablement d'enrichir la connaissance médicale, en apportant des réponses à des questions ouvertes telle que la recherche du *poids sec* en hémodialyse. L'évolution récente des technologies de l'information et des communications a, d'ores et déjà, entamé le rapprochement des univers, au départ fort éloignés que sont l'informatique et la médecine.

Publications

Publications avec une audience en informatique

[1] Julien Thomas, Cédric Rose, and François Charpillet. A multi-HMM approach to ECG segmentation. In *18th IEEE International Conference on Tools with Artificial Intelligence - ICTAI'06*, pages 609 –616, nov. 2006.

[2] Abdallah Dib, Cédric Rose, and François Charpillet. Bayesian 3d human motion capture using factored particle filtering. In *Proceedings of the 22th IEEE International Conference on Tools with Artificial Intelligence*, 2010.

[3] Cédric Rose, Jamal Saboune, and François Charpillet. Reducing particle filtering complexity for 3d motion capture using dynamic bayesian networks. In *AAAI'08 : Proceedings of the 23rd national conference on Artificial intelligence*, pages 1396–1401. AAAI Press, 2008.

[4] Cédric Rose, Chérif Smaili, and François Charpillet. A Dynamic Bayesian Network for Handling Uncertainty in a Decision Support System Adapted to the Monitoring of Patients Treated by Hemodialysis. In *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence - ICTAI'05*, Hong Kong/China, nov. 2005.

[5] Cédric Rose and François Charpillet. Apprentissage d'une discrétisation pour construire une politique à partir d'exemples. In *Acte des Journées Francophones Planification Décision Apprentissage (JFPDA 2009)*, PARIS France, juin 2009.

[6] Cherif Smaili, François Charpillet, Maan El Badaoui El Najjar, and Cédric Rose. Multi-sensor fusion for mono and multi-vehicle localization using Bayesian network. In Paula Fritzsche, editor, *Tools in Artificial Intelligence*. IN-TEH, 2008.

Publications avec une audience dans le domaine de la santé

[7] Jamal Saboune, Cédric Rose, and François Charpillet. Factored Interval Particle Filtering for Gait Analysis. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBC'07*, Lyon France, 2007. IEEE.

[8] Julien Thomas, Cédric Rose, and François Charpillet. A Support System for ECG Segmentation Based on Hidden Markov Models. In *Proceedings of the 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society - EMBC'07*, Lyon France, 2007. IEEE.

[9] Chérif Smaili, Cédric Rose, and François Charpillet. A decision support system for the monitoring of patients treated by hemodialysis based on a bayesian network. In *Pro-*

ceedings of the 3rd European Medical & Biological Engineering Conference - EMBEC'05, Prague/République Tchèque, nov. 2005.

[10] Jacques Chanliau, Cédric Rose, Luis Vega, and François Charpillet. Apport d'un système expert dans la définition du poids sec. In *Actes du 35ème Séminaire d'Uro-Néphrologie*, Paris, France, janv. 2009.

[11] Cédric Rose, Bernard Béné, François Charpillet, and Jacques Chanliau. Automatic Evaluation of Vascular Access in Hemodialysis Patients. In *Proceedings of the XXXVII European Society for Artificial Organs Congress - ESAO'10*, Skopje, R. Macedonia, sept. 2010.

Références

[Auvinet *et al.*, 2003] B. Auvinet, G. Berrut, C. Touzard, and L. Moutel. Gait abnormalities in elderly fallers. *Journal of aging and physical activity*, 2003.

[Baird, 1995] Leemon Baird. Residual algorithms : Reinforcement learning with function approximation. In *Proceedings of the Twelfth International Conference on Machine Learning*, pages 30–37. Morgan Kaufmann, 1995.

[Barrow *et al.*, 1977] H. G. Barrow, J. M. Tenenbaum, R. C. Bolles, and H. C. Wolf. Parametric correspondence and chamfer matching : two new techniques for image matching. In *IJCAI'77 : Proceedings of the 5th international joint conference on Artificial intelligence*, pages 659–663, San Francisco, CA, USA, 1977. Morgan Kaufmann Publishers Inc.

[Baum *et al.*, 1970] Leonard E. Baum, Ted Petrie, George Soules, and Norman Weiss. A maximization technique occurring in the statistical analysis of probabilistic functions of Markov chains. *The Annals of Mathematical Statistics*, 41(1) :164–171, 1970.

[Bertsekas and Tsitsiklis, 1996] Dimitri P. Bertsekas and John N. Tsitsiklis. *Neuro-Dynamic Programming*. Athena Scientific, 1996.

[Bradtke *et al.*, 1996] Steven J. Bradtke, Andrew G. Barto, and Pack Kaelbling. Linear least-squares algorithms for temporal difference learning. In *Machine Learning*, pages 22–33, 1996.

[Butler, 2003] Matthew Butler. Hidden Markov model clustering of acoustic data, 2003.

[Cassandra, 1998] Anthony Rocco Cassandra. *Exact and approximate algorithms for partially observable markov decision processes*. PhD thesis, Brown University, Providence, RI, USA, 1998. Adviser-Kaelbling, Leslie Pack.

[Chrisman, 1992] Lonnie Chrisman. Reinforcement learning with perceptual aliasing : The perceptual distinctions approach. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, pages 183–188. AAAI Press, 1992.

[Coulom, 2002] Rémi Coulom. *Reinforcement Learning Using Neural Networks, with Applications to Motor Control*. PhD thesis, Institut National Polytechnique de Grenoble, 2002.

[Cowell *et al.*, 1999] Robert G. Cowell, Steffen L. Lauritzen, A. Philip David, and David J. Spiegelhalter. *Probabilistic Networks and Expert Systems*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 1999.

[Deutscher *et al.*, 2000] J. Deutscher, A. Blake, and I. Reid. Articulated body motion capture by annealed particle filtering. In *Computer Vision and Pattern Recognition*, volume 2, pages 126–133, Hilton Head Island, SC, USA, 2000.

[Dubois, 2010] Amandine Dubois. *Les NTIC pour le maintien à domicile des personnes âgées*. PhD thesis, Université Nancy 2, 2010.

[Engel *et al.*, 2005] Yaakov Engel, Shie Mannor, and Ron Meir. Reinforcement learning with gaussian processes. In *In Proc. of the 22nd International Conference on Machine Learning*, pages 201–208. ACM Press, 2005.

[Ernst *et al.*, 2005] Damien Ernst, Pierre Geurts, Louis Wehenkel, and L. Littman. Tree-based batch mode reinforcement learning. *Journal of Machine Learning Research*, 6 :503–556, 2005.

[Falkhausen *et al.*, 1995] Markus Falkhausen, Herbert Reininger, and Dietrich Wolf. Calculation of distance measures between hidden markov models. In *Proceedings of Eurospeech*, pages 1487–1490, 1995.

[Fine and Singer, 1998] Shai Fine and Yoram Singer. The hierarchical hidden markov model : Analysis and applications. In *MACHINE LEARNING*, pages 41–62, 1998.

[Fung and Chang, 1990] Robert M. Fung and Kuo-Chu Chang. Weighing and integrating evidence for stochastic simulation in Bayesian networks. In *UAI '89 : Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence*, pages 209–220, Amsterdam, The Netherlands, The Netherlands, 1990. North-Holland Publishing Co.

[Gavrila, 1996] Dariu Mihai Gavrila. *Vision-based 3-D tracking of humans in action.* PhD thesis, University of Maryland, 1996.

[Geist and Pietquin, 2010] Matthieu Geist and Olivier Pietquin. A brief survey of parametric value function approximation. Technical report, Supélec, sept. 2010.

[Geist *et al.*, 2009] Matthieu Geist, Olivier Pietquin, and Gabriel Fricout. Kalman Temporal Differences : the deterministic case. In *IEEE International Symposium on Adaptive Dynamic Programming and Reinforcement Learning ADPRL 2009*, pages 185–192, Nashville, TN United States, 04 2009.

[Graja and Boucher, 2003] S. Graja and J.-M. Boucher. Markov models for automated ECG interval analysis. In *WISP 2003 : IEEE International Symposium on Inteligent Signal Processing*, pages 105–109, 2003.

[Heckerman and Shortliffe, 1992] David E. Heckerman and Edward H. Shortliffe. From certainty factors to belief networks. In *Artificial Intelligence in Medicine 4 :35–52*, 1992.

[Huang and Darwiche, 1996] C. Huang and A. Darwiche. Inference in belief networks : A procedural guide. *International Journal of Approximate Reasoning*, 15(3) :225–263, 1996.

[Hughes *et al.*, 2004] N. P. Hughes, L. Tarassenko, and S. J. Roberts. Markov models for automated ECG interval analysis. In S. Thrun, L. Saul, and B. Schoelkopf, editors, *Advances in Neural Information Processing Systems 16*, Cambridge, MA, 2004. MIT Press.

[Isard and Blake, 1998] Michael Isard and Andrew Blake. Condensation – conditional density propagation for visual tracking. *International Journal of Computer Vision*, 29(1) :5–28, 1998.

[Jain *et al.*, 1999] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering : a review. *ACM Comput. Surv.*, 31(3) :264–323, 1999.

[Jeanpierre, 2002] Laurent Jeanpierre. *Apprentissage et adaptation pour la modélisation stochastique de systèmes dynamiques réels.* PhD thesis, Université Henri Poincaré - Nancy I, 12 2002.

[Jensen *et al.*, 1990Pe] F. V. Jensen, S. L. Lauritzen, and K. G. Olesen. Bayesian updating in causal probabilistic networks by local computations. *Computational Statistics Quaterly*, 4 :269–282, 1990Pe.

[Jordan, 1998] Michael I. Jordan, editor. *Learning in Graphical Models.* MIT Press, 1998.

[Juang and Rabiner, 1985] B.-H. Juang and L. Rabiner. A probabilistic distance measure for hidden markov models. *AT&T Technical Journal*, pages 391–408, feb. 1985.

[Khawaja *et al.*, 2005] A. Khawaja, S. Sanyal, and O Dössel. A wavelet-based multi-channel ECG delineator. In *Proceedings of The 3rd European Medical and Biological Engineering Conference*, 2005.

[Kičmerová *et al.*, 2005] D. Kičmerová, I. Provazník, and J. Bardoňová. Classification of ECG signals using wavelet transform and hidden Markov models. In *Proceedings of The 3rd European Medical and Biological Engineering Conference*, 2005.

[Köhler *et al.*, 2002] B.-U. Köhler, C. Hennig, and R. Orglmeister. The principles of software QRS detection. *Engineering in Medicine and Biology Magazine, IEEE*, 2002.

[Krogh *et al.*, 1994] Anders Krogh, Michael Brown, I. Saira Mian, Kimmen Sjölander, and David Haussler. Hidden Markov models in computational biology : applications to protein modeling. *Journal of Molecular Biology*, 235 :1501–1531, 1994.

[Lauritzen and Spiegelhalter, 1988] S. L. Lauritzen and D. J. Spiegelhalter. Local computations with probabilities on graphical structures and their application to expert systems. *Journal of the Royal Statistical Society. Series B (Methodological)*, 50(2) :157–224, 1988.

[Lauritzen and Wermuth, 1989] S. L. Lauritzen and N. Wermuth. Graphical models for associations between variables, some of which are qualitative and some quantitative. *The Annals of Statistics*, 17(1) :31–57, 1989.

[Lauritzen, 1992] Steffen L. Lauritzen. Propagation of probabilities, means and variances in mixed graphical association models. *Journal of the American Statistical Association*, 87 :1098–1108, 1992.

[Li and Biswas, 2000] C. Li and G. Biswas. A bayesian approach to temporal data clustering using hidden markov models. In *Proceedings of the Seventeenth International Conference on Machine Learning*, pages 543–550, 2000.

[Littman *et al.*, 1995] Michael L. Littman, Anthony R. Cassandra, and Leslie Pack Kaelbling. Learning policies for partially observable environments : Scaling up. In *Proceedings of the Twelfth International Conference on Machine Learning*, pages 362–370. Morgan Kaufmann, 1995.

[Mannor *et al.*, 2004b] Shie Mannor, Duncan Simester, Peng Sun, and John N. Tsitsiklis. Bias and variance in value function estimation. In *Proc. of the 21st International Conference on Machine Learning*, pages 308–322, 2004.

[Martinez *et al.*, 2004] J. P. Martinez, R. Almeida, S. Olmos, A. P. Rocha, and P. Laguna. A wavelet-based ECG delineator : Evaluation on standard databases. In *IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING BME*, 2004.

[McCallum, 1992] R. A McCallum. First results with utile distinction memory for reinforcement learning. Technical report, University of Rochester, Rochester, NY, USA, 1992.

[McCormick and Isard, 2000] John McCormick and Michael Isard. Partitioned sampling, articulated objects, and interface-quality hand tracking. In *ECCV '00 : Proceedings of the 6th European Conference on Computer Vision-Part II*, pages 3–19, London, UK, 2000. Springer-Verlag.

[Meyer *et al.*, 2009] N. Meyer, S. Vinzio, and B. Goichot. La statistique bayésienne : une approche des statistiques adaptée à la clinique. *La revue de médecine interne*, 30(3) :242–249, 2009.

[Murphy, 1999] Kevin Murphy. Pearl's algorithm and multiplexer nodes. Technical report, 1999.

[Murphy, 2002] Kevin Patrick Murphy. *Dynamic Bayesian Networks : Representation, Inference and Learning.* PhD thesis, University of California Berkeley, 2002.

[Nedić and Bertsekas, 2002] A. Nedić and D. P. Bertsekas. Least squares policy evaluation algorithms with linear function approximation. *Theory and Applications*, 13 :79–110, 2002.

[Paysant, 2006] Jean Paysant. *Processus d'adaptation de la marche : l'apprentissage éco-contraint, influences des conditions de terrain sur la marche de l'amputé tibial.* PhD thesis, Université de Bourgogne, 2006.

[Pearl, 1988] Judea Pearl. *Probabilistic Reasoning in Intelligent Systems : Networks of Plausible Inference.* Morgan Kaufmann, 1988.

[Rabiner, 1989] L. R. Rabiner. A tutorial on hidden Markov models and selected applications in speech recognition. In *Proceedings of the IEEE*, pages 257–286, 1989.

[Saboune and Charpillet, 2005] J. Saboune and F. Charpillet. Using interval particle filtering for markerless 3d human motion capture. In *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence*, pages 621–627, Hong Kong, 2005.

[Saboune, 2008] Jamal Saboune. *Développement d'un système passif de suivi 3D du mouvement humain par filtrage particulaire.* PhD thesis, Université de technologie de Troyes, 2008.

[Schwarz, 1978] Gideon Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2) :461–464, 1978.

[Shachter and Andersen, 1994] Ross D. Shachter and Stig. K. Andersen. Global conditioning for probabilistic inference in belief networks. In *Proceedings of the Tenth Conference on Uncertainty in AI*, pages 514–522. Morgan Kaufmann, 1994.

[Shachter and Peot, 1990] Ross D. Shachter and Mark A. Peot. Simulation approaches to general probabilistic inference on belief networks. In *UAI '89 : Proceedings of the Fifth Annual Conference on Uncertainty in Artificial Intelligence*, pages 221–234, Amsterdam, The Netherlands, 1990. North-Holland Publishing Co.

[Shachter, 1998] Ross D. Shachter. Bayes-ball : The rational pastime (for determining irrelevance and requisite information in belief networks and influence diagrams). In *Proceedings of the Fourteenth Conference in Uncertainty in Artificial Intelligence*, pages 480–487, 1998.

[Shortliffe and Buchanan, 1975] Edward H. Shortliffe and Bruce G. Buchanan. A model of inexact reasoning in medicine. *Mathematical Biosciences*, 23(3-4) :351 – 379, 1975.

[Simon and Acker, 2008] Pierre Simon and Dominique Acker. La place de la télémédecine dans l'organisation des soins. Technical report, Ministère de la Santé et des Sports, Direction de l'Hospitalisation et de l'Organisation des Soins, nov. 2008.

[Smart, 2002] William Donald Smart. *Making reinforcement learning work on real robots.* PhD thesis, Brown University, Providence, RI, USA, 2002. Adviser-Kaelbling, Leslie Pack.

[Smyth, 1997] Padhraic Smyth. Clustering sequences with hidden Markov models. In *Advances in Neural Information Processing Systems*, pages 648–654. MIT Press, 1997.

[Sobel and Feldman, 1968] I. Sobel and G. Feldman. A 3x3 isotropic gradient operator for image processing. Technical report, talk presented at the Stanford Artificial Project, 1968.

[Sutton *et al.*, 2009] Richard S. Sutton, Hamid Reza Maei, Doina Precup, Shalabh Bhatnagar, David Silver, Csaba Szepesvári, and Eric Wiewiora. Fast gradient-descent methods for temporal-difference learning with linear function approximation. In *Proceedings of the 26th International Conference on Machine Learning*, 2009.

[Watkins and Dayan, 1992] Chris Watkins and Peter Dayan. Q-learning. *Machine Learning*, 8(3) :279–292, 1992.

[Yu and Bertsekas, 2007] Huizhen Yu and Dimitri P. Bertsekas. Q-learning algorithms for optimal stopping based on least squares. In *Proceedings of European Control Conference*, 2007.

[Yu *et al.*, 1979] Victor L. Yu, Lawrence M. Fagan, Sharon M. Wraith, William J. Clancey, A. Carlisle Scott, John Hannigan, Robert L. Blum, Bruce G. Buchanan, and Stanley N. Cohen. Antimicrobial selection by a computer, a blinded evaluation by infectious diseases experts. *The Journal of the Americam Medical Association*, 242(12), 1979.

Résumé

La télémédecine est une approche nouvelle de la pratique médicale qui est particulièrement porteuse d'espoir face à l'enjeu sociétal posé par l'incidence croissante des maladies chroniques et l'évolution de la démographie médicale. Le développement de la télésurveillance médicale réalisée grâce au recueil de données physiologiques ou biologiques au domicile du patient implique de développer nos capacités à analyser un volume important de données. Le problème auquel s'intéresse cette thèse est d'établir ou d'apprendre automatiquement la fonction qui lie les données fournies par les capteurs à l'état de santé du patient. La difficulté principale tient à ce qu'il est difficile et souvent impossible d'établir de manière sûre l'état de santé d'un patient, la seule référence disponible étant alors celle que peut donner le médecin traitant.

Nous montrons dans cette thèse que la modélisation stochastique et plus particulièrement le formalisme graphique bayésien permet d'aborder cette question sous trois angles complémentaires. Le premier est celui de la représentation explicite de l'expertise médicale. Cette approche est adaptée aux situations dans lesquelles les données ne sont pas accessibles et où il est donc nécessaire de modéliser directement la démarche du médecin. La seconde approche envisagée est celle de l'apprentissage automatique des paramètres du modèles lorsque suffisamment de données sur les sorties attendues sont disponibles. Nous nous intéressons enfin à la possibilité d'apprendre les actions pertinentes par renforcement sous les contraintes de la problématique médicale à savoir d'après l'observation de l'expert dans sa pratique normale. Nous étudions plus spécifiquement l'utilisation de la vraisemblance du modèle pour apprendre une représentation pertinente de l'espace d'états.

Abstract

Telemedicine is a new approach of medical practice that is expected to be one of the answers for facing the challenge of chronic diseases management. Development of remote medical surveillance at home relies on our capacity to interpret a growing amount of collected data. In this thesis, we are interested in defining the function that connects the state of the patient to the data given by the different sensors. The main difficulty comes from the uncertainty when assessing the state of the patient. The only reference available is the one that can be given by the medical doctor.

We show in this thesis that stochastic modelling and more specifically graphical bayesian formalism allows to treat this question in three ways. The first one consists in representing explicitly the medical expertise. This approach is adapted to the cases in which data is not accessible, and as a consequence, where it is necessary to model directly the diagnosis rules. The second approach that we study is the automatic learning of model parameters that can be performed when enough information is available concerning the expected outputs of the system. Finally, we propose the use of reinforcement for learning medical actions from the observation of the human expert in its everyday practice. Considering the specificity of the medical domain, we study the likelihood criterion for learning an efficient representation of the state space.